

Grundlagenstudien aus Kybernetik und Geisteswissenschaft

Akademia Libroservo/IfK
Kleinenberger Weg 16B
D-33100 Paderborn

Die Humankybernetik (Anthropokybernetik) umfaßt alle jene Wissenschaftszweige, welche nach dem Vorbild der neuzeitlichen Naturwissenschaft versuchen, Gegenstände, die bisher ausschließlich mit geisteswissenschaftlichen Methoden bearbeitet wurden, auf Modelle abzubilden und mathematisch zu analysieren. Zu den Zweigen der Humankybernetik gehören vor allem die Informationspsychologie (einschließlich der Kognitionsforschung, der Theorie über „künstliche Intelligenz" und der modellierenden Psychopathometrie und Geriatrie), die Informationsästhetik und die kybernetische Pädagogik, aber auch die Sprachkybernetik (einschließlich der Textstatistik, der mathematischen Linguistik und der konstruktiven Interlinguistik) sowie die Wirtschafts-, Sozial- und Rechtskybernetik. - Neben diesem ihrem hauptsächlichen Themenbereich pflegen die GrKG/Humankybernetik durch gelegentliche Übersichtsbeiträge und interdisziplinär interessierende Originalarbeiten auch die drei anderen Bereiche der kybernetischen Wissenschaft: die Biokybernetik, die Ingenieurkybernetik und die Allgemeine Kybernetik (Strukturtheorie informationeller Gegenstände). Nicht zuletzt wird auch metakybernetischen Themen Raum gegeben: nicht nur der Philosophie und Geschichte der Kybernetik, sondern auch der auf kybernetische Inhalte bezogenen Pädagogik und Literaturwissenschaft. -

La prihoma kibernetiko (antropokibernetiko) inkluzivas ĉiujn tiajn sciencobranĉojn, kiuj imitante la novepokan natursciencon, klopodas bildigi per modeloj kaj analizi matematike objektojn ĝis nun pritraktitajn ekskluzive per kultursciencaj metodoj. Apartenas al la branĉaro de la antropokibernetiko ĉefe la kibernetika psikologio (inkluzive la ekkon-esploron, la teoriojn pri „artefarita intelekto" kaj la modeligajn psikopatometrion kaj geriatrion), la kibernetika estetiko kaj la kibernetika pedagogio, sed ankaŭ la lingvokibernetiko (inkluzive la tekststatistikon, la matematikan lingvistikon kaj la konstruan interlingvistikon) same kiel la kibernetika ekonomio, la socikibernetiko kaj la jurkibernetiko. - Krom tiu ĉi sia ĉefa temaro per superrigardaj artikoloj kaj interfake interesigaj originalaj laboraĵoj GrKG/HUMANKYBERNETIK flegas okaze ankaŭ la tri aliajn kampojn de la kibernetika scienco: la biokibernetikon, la inĝenierkibernetikon kaj la ĝeneralan kibernetikon (strukturteorion de informecaj objektoj). Ne lastavice trovas lokon ankaŭ metakibernetikaj temoj: ne nur la filozofio kaj historio de la kibernetiko, sed ankaŭ la pedagogio kaj literaturscienco de kibernetikaj sciaĵoj. -

Cybernetics of Social Systems comprises all those branches of science which apply mathematical models and methods of analysis to matters which had previously been the exclusive domain of the humanities. Above all this includes *information psychology* (including theories of cognition and 'artificial intelligence' as well as psychopathometrics and geriatrics), *aesthetics of information* and *cybernetic educational theory*, *cybernetic linguistics* (including text-statistics, mathematical linguistics and constructive interlinguistics) as well as *economic*, *social* and *juridical cybernetics*. - In addition to its principal areas of interest, the GrKG/HUMANKYBERNETIK offers a forum for the publication of articles of a general nature in three other fields: *biocybernetics*, *cybernetic engineering* and *general cybernetics* (theory of informational structure). There is also room for *metacybernetic* subjects: not just the history and philosophy of cybernetics but also cybernetic approaches to education and literature are welcome.

La cybernétique sociale contient tous le branches scientifiques, qui cherchent à imiter les sciences naturelles modernes en projetant sur des modèles et en analysant de manière mathématique des objets, qui étaient traités auparavant exclusivement par des méthodes des sciences culturelles („idéographiques"). Parmi les branches de la cybernétique sociale il y a en premier lieu la psychologie informationelle (inclues la recherche de la cognition, les théories de l'intélligence artificielle et la psychopathométrie et gériatrie modeliste), l'esthétique informationelle et la pédagogie cybernétique, mais aussi la cybernétique linguistique (inclues la statistique de textes, la linguistique mathématique et l'interlinguistique constructive) ainsi que la cybernétique en économie, sociologie et jurisprudence. En plus de ces principaux centres d'intérêt la revue GrKG/HUMANKYBERNETIK s'occupe - par quelques articles de synthèse et des travaux originaux d'intérêt interdisciplinaire - également des trois autres champs de la science cybernétique: la biocybernétique, la cybernétique de l'ingenieur et la cybernétique générale (théorie des structures des objets informationels). Une place est également accordée aux sujets métacybernétiques mineurs: la philosophie et l'histoire de la cybernétique mais aussi la pédagogie dans la mesure où elle concernent la cybernétique.

ISSN 0723-4899

Grundlagenstudien aus Kybernetik und Geisteswissenschaft

Internationale Zeitschrift für Modellierung und Mathematisierung in den Humanwissenschaften
Internacia Revuo por Modeligo kaj Matematikizo en la Homsciencoj
International Review for Modelling and Application of Mathematics in Humanities
Revue internationale pour l'application des modèles et de la mathématique en sciences humaines



HUMANKYBERNETIK

Inhalt * Enhavo * Contents * Sommaire

Band 38 * Heft 3 * Sept. 1997



Akademia Libroservo

Die GRUNDLAGENSTUDIEN AUS KYBERNETIK UND GEISTESWISSENSCHAFT

(grkg/Humankybernetik) wurden 1960 durch Max BENSE, Gerhard EICHHORN und Helmar FRANK begründet. Sie sind z.Zt. offizielles Organ folgender wissenschaftlicher Einrichtungen:

INSTITUT FÜR KYBERNETIK BERLIN e.V.
Gesellschaft für Kommunikationskybernetik
(Direktor: Prof.Dr.phil.habil. Heinz Lohse, Leipzig, D)

TAKIS - Tutmonda Asocio pri Kibernetiko, Informadiko kaj Sistemiko
(prezidanto: D-ro Dan Maxwell, Washington, USA; ĝenerala sekretario: Inĝ. Milan Zvara, Poprad, SK)

AKADEMIO INTERNACIA DE LA SCIENCOJ (AIS) San Marino
publikigadas siajn oficialajn sciigojn komplete en grkg/Humankybernetik

Grundlagenstudien aus Kybernetik und Geisteswissenschaft

Internationale Zeitschrift für Modellierung und Mathematisierung in den Humanwissenschaften
Internacia Revuo por Modeligo kaj Matematikizo en la Homsciencoj
International Review for Modelling and Application of Mathematics in Humanities
Revue internationale pour l'application des modèles et de la mathématique en sciences humaines



Inhalt * Enhavo * Contents * Sommaire

Band 38 * Heft 3 * Sept. 1997



Akademia Libroservo

Verlag und Anzeigenverwaltung **Eldonejo kaj anoncadministrejo** **Publisher and advertisement administrator** **Edition et administration des annonces**



Akademia Libroservo - Internacia Eldongrupo Scienca:
AIEP - San Marino, Esprima - Bratislava, Kava-Pech - Dobrichovice/Praha,
IfK GmbH - Berlin & Paderborn, Libro - Jelenia Góra
Gesamtherstellung: **IfK GmbH**
Verlagsabteilung: Kleinenberger Weg 16 B, D-33100 Paderborn,
Telefon (0049-/0-)5251-64200 Telefax: -163533

Die Zeitschrift erscheint vierteljährlich (März, Juni, September, Dezember). Redaktionsschluß: 1. des vorigen Monats. - Die Bezugsdauer verlängert sich jeweils um ein Jahr, wenn bis zum 1. Dezember keine Abbestellung vorliegt. - Die Zusendung von Manuskripten (gemäß den Richtlinien auf der dritten Umschlagseite) wird an die Schriftleitung erbeten, Bestellungen und Anzeigenaufträge an den Verlag. - Z. Zt. gültige Anzeigenpreisliste auf Anforderung.

La revuo aperadas kvaronjare (marte, junie, septembre, decembre). Redakcia limdato: la 1-a de la antaŭa monato. - La abondaŭro plilongiĝas je unu jaro se ne alvenas malmendo ĝis la unua de decembro. - Bv. sendi manuskriptojn (laŭ la direktivoj sur la tria kovrilpaĝo) al la redakcio, mendojn kaj anoncojn al la eldonejo. - Momente valida anoncprezlisto estas laŭpete sendota.

This journal appears quarterly (every March, Juni, September and December). Editoial deadline is the 1st of the previous month. - The subscription is extended automatically for another year unless cancelled by the 1st of December. - Please send your manuscripts (fulfilling the conditions set our on the third cover page) to the editorial board, subscription orders and advertisements to the publisher. - Current prices for advertisements at request.

La revue est trimestrielle (parution en mars, juin, septembre et décembre). Date limite de la rédaction: le 1er du mois précédent. L'abonnement se prolonge chaque fois d'un an quand une lettre d'annulation n'est pas arrivée le 1er décembre au plus tard. - Veuillez envoyer, s.v.p., vos manuscrits (suivant les indications de l'avant-dernière page) à l'adresse de la rédaction, les abonnements et les demandes d'annonces à celle de l'édition. - Le tarif des annonces en vigueur est envoyé à la demande.

Bezugspreis: Einzelheft 20,-- DM; Jahresabonnement: 80,-- DM plus Versandkosten.

Druck: Druckerei Reike GmbH, D-33106 Paderborn

Denove pri la plej originala, la plej probabla kaj la plej akceptebla solvo de problemo

de Zdeněk PŮLPÁN, Hradec Králové (CZ)

el la Vysoká škola pedagogická Hradec Králové

En ĉi tiu artikolo mi provas denove difini gradon de la originaleco (vidu Půlpán 1993, Frank 1993), poste mi rememorigos la algoritmon de la plej verŝajna solvo kaj rimarkigos eblecon de trovo de la plej akceptebla solvo sur la bazo de la ekspertaj taksoj.

Ni supozas, ke ĉiuj solvoj de certa klaso de reprezentantaj taskoj povas esti priskribitaj per orientita arba grafikaĵo, kies nodoj prezentas decidadojn kaj strekoj prezentas certajn decidojn. Stimante probablojn de apartaj decidaĵoj, la probablo de certa solvo estas donata per produkto de probabloj de korespondantaj strekoj de ĝiaj partaj solvoj. Se k estas la nombro de ĉiuj eblaj solvoj kaj $P(i)$ la probablo de la i-a solvo ($1 \le i \le k$) la kvanto H de informacio pri elekto de certa tipo de solvo per hazarde elektita respondanto estas donata per entropio (1):

$$H = -\sum_{i=1}^{k} P(i) \cdot \log_2 P(i) \qquad [\text{bit}] \tag{1}$$

La solvo de tasko estas ĉi tie komprenata kiel fina progreso de partaj solvoj (en nia grafika interpretado tio korespondas al orientita vojo de radiko al certa folio).

En ĉiu parta solvita paŝo oni povas stimi transinformon I kiel la sumon

$$I = \log_2 n + \sum_i p_i \log_2 p_i ; \quad \sum_i p_i = 1,$$

kie n estas la nombro de eblaj decidaĵoj en tiu paŝo kun la probabloj $p_1, p_2, \ldots, p_n (\Sigma\, p_i = 1)$. Poste por ĉiu progreso de solvitaj paŝoj oni povas fiksi maksimumon el partaj transinformoj. Por individualaj solvoj oni tiel ricevos valorojn de respondditaj maksimumoj $I_1, I_2, \ldots I_k$. Ni proponos poste indiki tiun solvon kiel la plej originala (en nia senco), por kiu la korespondanta valoro de I estas la plej granda (el ĉiuj nodoj de apartenanta solvo) kaj al tio la apartenanta vojo enhavas strekon de nodo kun la plej granda valoro de I, kiu havas ankaŭ la plej malgrandan probablon el ĉiuj elirantaj strekoj de la nodo.

Mezuro de originaleco devas esti relativa, t.e. rilatanta al destinata referenca algoritmo. Tio estos ekzemple la algoritmo al kiu apartenas vojo kun la plej malgranda maksimumo I^0 de transinformacioj el ĉiuj nodoj kaj kun streko eliranta el tiu nodo kun la plej granda probablo p^0. Se la maksimumo de transinformacioj el ĉiuj nodoj de korespondanta algoritmo I estas la mezuro de originaleco ω egalas al

$$\omega = \frac{(1-p)\cdot I - (1-p^0)\cdot I^0}{(1-p^0).I^0}; \quad \omega \succ 0 \tag{2}$$

kie p estas la probablo korespondanta al elira streko el nodo kun la plej granda transinformacio, kiu estas parto de konsiderata algoritmo.

Pli grandaj valoroj de ω atestas pri granda originaleco de la solvo.

La plej probablan vojon en grafikaĵo oni trovos per konvena algoritmo (ekz. per metodo de Danzig). Ni povas procede al ĉiuj nodoj de grafikaĵo aldoni valoron v_i:

a) al enira nodo ni aldonos la valoron 1, al la aliaj nodoj ni fiksas valoron 0;

b) ni serĉos tiujn strekojn inter nodoj i, j, por kiuj

$$v_i \prec v_i \cdot p_{ij}$$

kie $p_{ij} \in \prec 0;1 \succ$ estas la probablo de transiro en grafikaĵo el la i-a nodo en la j-an nodon; se ni trovos tian strekon, ni ŝanĝos la valoron de la variablo v_j en

$$v'_j = v_i \cdot p_{ij};$$

c) laŭ b) ni procedos ĝis por ĉiuj strekoj ij valoros

$$v_j \geq v_i \cdot p_{ij} \tag{3}$$

d) Tuj kiam la procedo laŭ antaŭpaŝoj estas finita, estas eble difini tian progreson (aŭ tiajn progresojn) de strekoj, kiu (aŭ kiuj) kunigas eniron (en radikon) kun eliro (kun korespondanta folio), por kiuj

$$v_j = v_i \cdot p_{ij}. \tag{4}$$

Tuj strekoj kreas la optimuan vojon kun la maksimuma probablo egala al la valoro v de eliro.

Dum valorado de solvo de certa problemo estu la valoradisto kapabla siajn imagojn kaj postulojn por solvo antaŭe specifi. Kompleta specifo ne estas ebla. Tial por objektivigo de valorado estas necese respekti tiun metodon, kiu el pli multaj vidpunktoj argumentate elektos la plej relevantajn vidpunktojn. Tiuj vidpunktoj tamen povas esti konfliktaj aŭ komunaj, per ilia akceptado povas diversaj taksantoj gajni diversajn taksaĵojn de la solvo de problemo.

Ĉiu taksado en la situacioj kiam la taksanto estas homo, estas karakterizata per neprecizeco kaj kompromiso en multaj aspektoj. Tion necesos respekti dum interpretado de preferencoj de variantoj de solvo.

Ni marku relevantajn vidpunktojn de solvo de problemo per $R_1, R_2 \ldots, R_k$ kaj per $x_1, x_2, \ldots, x_e$ variantojn de eblaj solvoj de problemo.

Ni supozu, ke al ĉiu vidpunkto de taksado R_i estas donata sinsekvo de „konveneco" de variantoj (la plej konvena varianto havos la plej malaltan sinsekvon).

Unu el kriterioj por fikso de variantoj estas kalkulo de sumo de iliaj sinsekvoj por apartaj vidpunktoj. Plej bone estos poste taksata tiu varianto, kiu havos la plej malaltan sumon de sinsekvoj. Sed oni supozas, ke ĉiuj sinsekvoj kaj ĉiuj karakterizaĵoj havas egalan pezon.

Tial estas iam konvene enkonduki matricon $A = (a_{ij})_{e,e}$, kies elementoj a_{ij} determinas, kiomoble estas varianto x_i taksata kiel j-a en sinsekvo de ĉiuj variantoj.

Se estas donita vektoro $\vec{v}$ de pezoj de individuaj vidpunktoj $\vec{v} = (v_1, v_2, \ldots, v_k)$, $(v_i \geq 0; \sum v_i = 1)$, oni povas difini ankaŭ matricon $B = (b_{ij})_{e,e}$, kies elementoj b_{ij} estas la sumo de pezoj de tiuj karakterizaĵoj, laŭ kiuj estas varianto x_i taksata kiel en j-a en sinsekvo. Elementon b_{ij} oni povas poste interpreti kiel (nenormatan) mezuron de konveneco de aldonaĵo de la j-a sinsekvo al la i-a varianto.

Literaturo:

Černý, M., D. Glückaufová: Vícekriteriální rozhodování za neurčitosti, Academia, Praha 1987

Černý, M., D. Glückaufová, M. Toms: Metody komplexního vyhodnocování variant, Academia, Praha 1980

Dantzig, G., B.: On the Shortest Route Through a Network, Management Science, 6, Nr. 2

Ford, L., R., D., R. Fulkerson: Flows in Networks, Princeton Princeton University Press 1962

Frank, H.: Noto pri la PŮLPÁNa originalec-mezuro de problemsolvoj grkg/Humankybernetik, Band 34, Heft 4, 1993

Půlpán, Z., F. Kuřina, Vl. Kebza: O představivosti a její roli v matematice, Academia, Praha 1991

Půlpán, Z.: La plej konvena kaj la plej originala solvo, grkg/Humankybernetik, Band 34, Heft 4, 1993

Půlpán, Z.: Odhad míry originality myšlení, Psychológia a patopsychológia dieťaťa, 23, 1988, Nr. 3

Ricevita 1994-09-23

Adreso de la aŭtoro: Doc. RNDr. Zdeněk Půlpán, Karla IV-789, CZ-50002 Hradec Králové I

New about the most original, probable and acceptable problem solution (Summary)

In the research C - R tests of the predominantly mathematical way of thinking proved that significant characteristics of the mathematical way of thinking are the basis for the original solution made by the standard population. It concerns the analysis of single steps leading to the task solution when using a tree graph. Using the probability calculus of exact algorithms for the solution, it is possible to compare levels of various populations solving the same task.

The paper presents the method for estimation of the best original solution as well as the most probable and preferable when using external a priori statements.

grkg / Humankybernetik
Band 38 · Heft 3 (1997)
Akademia Libroservo / IfK

Versuch einer ergänzenden Auswertung des Schlüsselversuchs der Informationspsychologie

von Helmar FRANK, Paderborn (D), und Diana FALOBA, Sibiu/Hermannstadt (RO)

aus dem Institut für Kybernetik / Universität Paderborn (Direktor: Prof.Dr. Helmar Frank)

1. *„Bewußtsein" als Schlüsselbegriff der Kybernetik*

„Man kann ohne den Begriff des Bewußtseins nicht in die Kybernetik hineinkommen, mit ihm aber nicht in ihr bleiben" wurde an anderer Stelle (Frank, 1975, S. 51) zur Kennzeichnung des Konflikts zwischen Thema und Ziel der Kybernetik formuliert. Der programmatische Aufbau der Kybernetik macht das „empfangene Zeichen" zum Elementarbegriff, dessen Bestimmung „nur *relativ zu einem Subjekt* möglich" ist, welches seinerseits „zunächst naiv als *Bewußtsein* verstanden werden kann" (Frank, 1995, S. 65; 1996a, S. 47f.). *Ziel* ist aber die (im Sinne von Schmidt, 1941:) *Objektivierung* geistiger Arbeit, zumindest die *Vorausberechenbarkeit mentaler Prozesse* im Sinne des für nomothetische Wissenschaften (im Sinne von W. Windelband) geltenden Comteschen Imperativs „savoir pour prévoir". Beide Ziele machen *methodisch* eine modellierende Rekonstruktion des vorgegebenen und vorab cartesisch analysierten Phänomens erforderlich, damit Modellparameter meßbar und miteinander verrechenbar werden (vgl. den nomothetischen Methodenfünfschritt in Frank, 1995, S. 62 - 65; 1996a, S. 41 - 44). Das Subjekt wird dabei in der Objektwelt nachgebildet, wobei die ursprüngliche *Ichigkeit* und *Einheit* des Bewußtseins, also seine Wesensmerkmale, außer Betracht bleiben. Stattdessen gilt es, partielle Fähigkeiten, wie insbesondere die Grenzen für die Schnelle der Aufnahme sowie für Dauer und Umfang der Wahrung empfangener Zeichen, meßbar zu machen.

Zur Spezialisierung auf die *Human*kybernetik, gehört es, die - in erster Linie durch die mathematische Informationstheorie Shannons (1948) - *allgemein* erreichte Meß*barkeit* von Informationellem *konkret anzuwenden*, und zwar nicht, wie in der *Konstrukt*kybernetik (d.h. der Ingenieur- oder Maschinenkybernetik), zur Aufstellung von Vorschriften für eine zu konstruierende (maschinentechnische) Objektivation, sondern zur quantitativen Beschreibung und exakten Erklärung der im ursprünglichen Subjekt selbst empirisch schon vorfindlichen, mentalen Prozesse. Dies gilt insbesondere für die Mikrohumankybernetik (Informationspsychologie und -ästhetik, Bildungs- und Sprachkkybernetik), die an der individuellen Informationsverarbeitung interessiert ist und auch Kommunikations-

kybernetik genannt wird (Frank, 1993, S. 192; Frank/Franke, 1997, S. 55 - 57). Dabei war zunächst die fundamentale Frage zu beantworten, ob der Mensch bei seiner mentalen Aktivität überhaupt irgendwo (unbewußt) ausschöpft, was informationstheoretisch möglich ist, ob also die Shannonsche Informationstheorie für die Humankybernetik überhaupt *hilfreich* ist.

Nach Shannons Fundamentalsatz gibt es zu jedem Informationsübertragungskanal (zu einer anderen Stelle oder - bei Speichern - bis zu einem späteren Zeitpunkt) eine unüberschreitbare Grenze seiner Übertragungsfähigkeit; genannt *Kanalkapazität C,* die aber bei geeigneter Codierung beliebig gut genutzt werden kann. Dabei sind einerseits C, andererseits die zu übertragende Nachricht in der von Shannon eingeführten und (auf Anregung von J.W. Tukey) „bit" genannten Informationseinheit zu messen, und die *Übertragbarkeit* einer Nachricht *nach ihrer geeigneten Codierung* besteht genau dann, wenn ihr so gemessener (zu speichernder oder pro Zeiteinheit zu übertragender) Informationsgehalt die Kanalkapazität C nicht übersteigt. *Ob im Einzelfall* Nachrichten in diesem, durch C begrenzten Höchstumfang übertragen, also *tatsächlich* geeignet codiert werden, bleibt dabei offen und muß in jedem konstrukt-, bio- oder humankybernetischen Anwendungsfall untersucht werden; Shannons Informationstheorie gehört zur *allgemeinen* Kybernetik.

2. *Der Schlüsselversuch zur Aufnahmefähigkeit des menschlichen Bewußtseins.*

Die Informationstheorie kann insbesondere für die Psychologie nur dann hohe Relevanz haben, wenn sich feststellen läßt, daß

a) von Zeichenketten gleicher Länge in gleicher Zeit *mehr* Zeichen wahrgenommen werden, wenn sie weniger „Information" enthalten (im Sinne des Shannonschen Informationsmaßes; Hilgers und Yashovardhan [1980, S. 37] führten dafür in ILo den Neologismus *informacio* ein), aber

b) von diesen Zeichenketten in dieser Zeit *gleichviel Informacio* wahrgenommen wird, und

c) daß Wahrnehmungszeit und wahrgenommene Informacio proportional zueinander sind.

Für die Signalübertragung von den Rezeptoren, z.B. der Retina, zu den sensorischen Projektionszentren (also für die Perzeption) gilt dies sicher nicht: das häufige ***e*** wird nicht schneller *perzipiert* als das seltene ***x***. Wenn überhaupt, dann kann die zeitsparende Codierung also nur anschließend, vor der Bewußtwerdung (Apperzeption), erfolgen. Das Psychostrukturmodell der Informationspsychologie (die früheste graphische Fassung findet sich in Frank, 1961, S. 19) postuliert für diese informationstheoretisch mögliche Optimalkodierung den später (Frank/Meder, 1971, S. 41) so genannten „Akkomodator" zur Eingabe in den (beliebig kurz denkbaren) Kanal zum Bewußtsein. Die neurophysiologische (oder zumindest Modell-) Entsprechung des Bewußtseins wurde (Frank, 1961, S. 19f.) „Kurzspeicher" genannt. Die (zunächst im Hinblick auf ästhetische Probleme gestellte) Grundfrage lautet: Wird für diesen Kanal, also für die *Apperzeption* als zweiter Phase der Wahrnehmung, die informationstheoretisch bestehende Möglichkeit der redundanzproportionalen Zeiteinsparung genutzt? Die zu bestimmende Kanalkapazität C ist

dann als *Apperzeptionsschnelle* in einem darauf aufzubauenden informationspsychologischen Kalkül verwendbar.

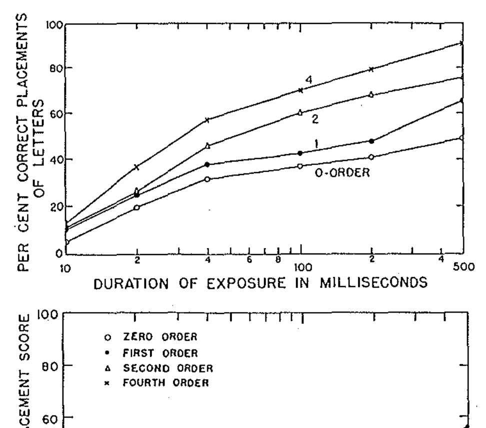


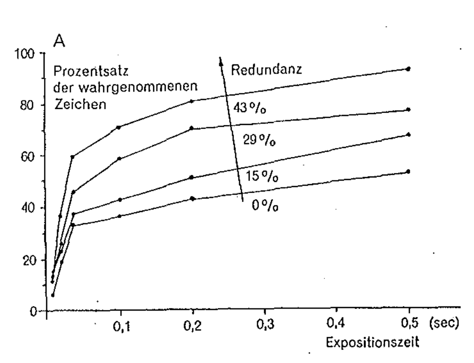
Bild 1: Von 10 - 500 msek lang gezeigten 8-Buchstaben-Folgen konnten Versuchspersonen einen größeren Prozentsatz P richtig und an der richtigen Stelle wiedergeben, wenn die Folgen gewohnte Texte besser annäherten (0. - 4. Ordnung), also redundanter waren (R = 0%; 15%; 29%; 43%), d. h. weniger Informacio insgesamt (I = 38 bit; 32 bit; 27 bit; 22 bit) bzw. pro Zeichen (H = 4,7; 4,0; 3,3; 2,7 bit) enthielten. (Aus Miller, Bruner, Postman, 1954, S. 135.)

Bild 2: Infolge gleichlanger Darbietungszeit wird redundanzunabhängig von jeder der in Bild 1 angesprochenen Buchstabenfolgen derselbe Betrag $IP = 8HP = 8(1\text{-}R)H_{max}P$ an Informacio aufgenommen, also derselbe Prozentsatz $IP/8H_{max} = IP/(38\text{ bit}) = (1\text{-}R)P$ der maximal in 8 Buchstaben unterbringbaren Informacio. (Aus Miller, Bruner, Postman, 1954, S. 138.)

Ohne daß dort schon dieser Modellansatz und damit die Unterscheidung zwischen Perzeption und Apperzeption gemacht wäre, kann der Versuch von George A. Miller, Jerome S. Bruner und Leo Postman (1954 [im folgenden kurz „M/B/P"] - die Arbeit wurde bereits im Januar 1952 zur Veröffentlichung eingereicht) als wahrscheinlich erste eindeutige Bestätigung der Relevanzbedingungen (a) - durch Bild 1 - und (b) - durch Bild 2 - gelten. Jedenfalls wurde er zum Schlüsselversuch, der zum Ansatz der Informationspsychologie führte (Frank, 1959, S. 28 - 31) und in deren Schrifttum als solcher ständig zitiert wird (ausführlich vor allem in Frank, 1962, S. 94 - 97, 1964, S. 82-84, 1969, Bd. 2, S. 78 - 82, 1996, S. 75 - 79; v. Cube, 1965, S. 88 - 93; Weltner, 1970, S. 112; Frank und Wagner, 1982, S. 73 - 76). Diesen Versuch nannte A. Moles Ende 1958 in Paris mit Verweis auf die kurz zuvor erschienene Darstellung durch Cherry (1957, S. 280 - 283) als Argument, als er (von Frank) kritisch im Sinne der oben aufgeführten Kriterien (a,b) nach der psychologischen Relevanz der Informationstheorie befragt wurde. - M/B/P selbst hatten (der Denkweise des US-amerikanischen Behaviorismus getreu) ihre Untersuchung *nicht* zur Bestimmung einer „Apperzeptionsschnelle" als Parameter für einen informationstheoretisch begründeten psychologischen Modellansatz ausgewertet.

Die Bedingung (c) ist eine Folge von (b) und der fast trivialen Unterstellung, daß - von Ermüdungserscheinungen abgesehen - von Zeichenketten konstanter mittlerer Informacio in jeder zusätzlichen Zeiteinheit gleichviel weitere Zeichen apperzipiert werden können. M/B/P machen hierzu keine Aussage. Die Originaldarstellung ihrer Ergebnisse (Bilder 1

und 2) läßt sogar vermuten, daß sie hieran überhaupt nicht dachten, denn sonst wäre die logarithmische Teilung der Zeitachse kaum noch durch die bessere Verteilung der Meßpunkte im Koordinatensystem zu rechtfertigen. Die Umzeichnung in lineare Koordinatensysteme (erstmals in Frank, 1959, S. 29; verbesserte Fassungen finden sich u.a. in Frank, 1962, S. 95, 1963, S. 83 und in v.Cube, 1965, S. 89) lag also nahe. Daß sie nicht zu einer augenfälligen Proportionalität zwischen der tachistoskopischen Expositionszeit und der wahrgenommenen Zeichenzahl oder Infomacio führte (Bild 3), wurde als Folge der „bekannten Nachwirkungserscheinungen" erklärt (Frank, 1959, S. 36f., 1962, S. 95f., 1969, Bd. 2, S. 80). Zeitproportionalität ist nur für die vom Akkomodator zum Kurzspeicher fließende Informacio zu erwarten. Proportionalität ist also nicht zur Expositionszeit (Reizzeit) zu vermuten, sondern zur Apperzeptionszeit; diese ist um eine Nachwirkungszeit (t_s) verlängert.



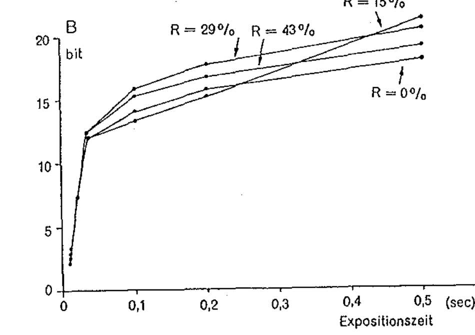


Bild 3: Die Umzeichnung von (A:) Bild 1 und (B:) Bild 2 in ein Koordinatensystem mit linearer statt logarithmischer Abszissenteilung führt statt zu einer Zeitproportionalität des Prozentsatzes wahrgenommenen Zeichen (A) bzw. der wahrgenommenen Informacio (B) zu einer Knickstelle zwischen 40 msek und 100 msek, also in der Größenordnung des SZQ. (Nach Frank, 1959, S. 29.)

Folgerichtig wurden daher schon in der informationsästhetischen Dissertation von Frank (1959, S. 29) als Apperzeptionsschnelle etwa 16 bit/sek dem Anstieg des zweiten Astes der M/B/P-Kurven entnommen. Dazu wurde keine genaue C-Bestimmung aus den Originaldaten von M/B/P versucht. Vielmehr wurde auf die mögliche Binärkodierung im Zentralnervensystem, auf die bei 1/16 sek liegende Schwelle des „Zeitsinns" (dem „Moment" oder „subjektiven Zeitquant [SZQ]) und auf die um 16 Hz streuenden elektroenzephalographischen Beta-Wellen vorab der Ansatz aufgebaut, die Apperzeptionsschnelle sei genau 1 bit/SZQ (Frank, 1959, S. 28) „oder allenfalls ein ganzzahliges Vielfaches davon" (Frank, 1962, S. 85). Der M/B/P-Versuch wurde somit zu einem *Bestätigungs*experiment für eine vorab aufzustellende Theorie umgewertet (oder allenfalls zur empirischen Entscheidung zwischen diskreten theoretischen Möglichkeiten für C).

Daß die zwischen 1/25 sek und 1/10 sek zu suchende Knickstelle der empirischen Kurven genau bei 1 SZQ ≈ 1/16 sek liegt, wurde zunächst ohne Begründung unterstellt

(Frank, 1959, S. 29, 1962, S. 95f., S. 98), später als „möglich" beiläufig erwähnt (Frank, 1969, Bd. 2, S. 80). -

Spätere Arbeiten führten zu weiteren Relativierungen, die es unzulässig machen, aus den M/B/P-Ergebnissen auf die natürliche Zahl n in der Gleichung $C = 16n$ bit/sek zu schließen. Riedel (1965) bestätigte die Vermutung von Frank (1962, S. 96f.), daß C, falls in bit/sek gemessen, bei Kindern altersabhängig wächst, und zugleich die in Sekunden gemessene Dauer des SZQ fällt, und zwar so, daß weiterhin (sozusagen als „Fundamentalsatz der Informationspsychologie")

$$(1)\ C = 1 \text{ bit/SZQ}$$

gilt. Lehrl (1981) stellte (gegen die Vermutung von Frank, 1962, S. 110, [2]1969, Bd. 2, S. 136f.) empirisch eine positive Korrelation zwischen dem IQ und der Apperzeptionsschnelle fest; Roth (1972, [4]1980, S. 95) hatte dies schon aus der negativen Korrelation zwischen IQ und Reaktionszeit in Verbindung mit Franks (1960; 1962, S. 122 - 125) Erklärung der Reaktionszeit als Summe der Apperzeptionszeit und der Zeit für die Assoziation semantischer Information geschlossen.

3. *Problemstellung*

Der geschichtliche Rückblick zeigt, daß der M/B/P-Versuch nicht nur Auslöser für die Begründung der Informationspsychologie wurde, sondern auch zu deren systematischer Grundlage, von welcher aus im Laufe der Jahre eine kohärente, zahlreiche empirische Einzelerkenntnisse umfassende Theorie aufgestellt werden konnte. (Ihre „Fortschrittsgeschwindigkeit" suchte E. Cistean [1995] zu messen.) Auch in der hochschulpädagogischen Aufarbeitung der Informationspsychologie spielt dieser Versuch die Rolle eines Schlüsselexperiments (vgl. z.B. Frank, 1969, Kapitel 5, und 1996, Kapitel 6). Insbesondere aus diesem Grunde bedürfen drei Mängel seiner bisherigen Auswertung einer Behebung:

1. Nur die Meßergebnisse für Expositionsdauern $d > 1$ SZQ wurden für die Theorie verwertet. Für $d < 1$ SZQ ist die Bedingung (b) augenscheinlich viel genauer erfüllt, was nicht einmal zu erklären versucht wurde.
2. Die mittlere Apperzeptionsgeschwindigkeit der benutzten Versuchspersonen (und - wie bei der Nachuntersuchung von Frank und Wagner, 1982, S. 78, geschehen - deren wahrscheinlicher IQ) kann und sollte empirisch bestimmt werden, ohne (1) schon vorauszusetzen.
3. Die Lage des Knickpunktes zwischen den je zwei Ästen der von M/B/P empirisch gefundenen Kurven sollte bestimmt werden, um festzustellen, ob die Unterstellung, er liege bei 1 SZQ, mit den Daten verträglich ist. Den „Fundamentalsatz" (1) unterstellend können dann die Meßwerte für $d < 1$ SZQ zu einer genaueren Bestimmung von Apperzeptionsgeschwindigkeit und SZQ-Dauer der Versuchspersonen mitgenutzt werden.

Die Mitautorin (Faloba, 1997) nahm diese ergänzende Auswertung für ihre Bakkalaureatsarbeit bei der Akademio Internacia de la Sciencoj (AIS) San Marino vor. Im folgenden werden die Ergebnisse knapp dargestellt.

4. *Reiznachwirkungszeit*

Schon von der Sinnesphysiologie (vgl. z.B. v. Buddenbrock, 1952, 1953; Rein/Schneider, 1955, u.a.) ist bekannt, daß die Reizung von Rezeptoren, z.B. der Retina, mit einer Latenzzeit t_1 die Sendung von Impulsen an das Gehirn auslöst, und daß diese Signale auch noch während einer Zeit $t_2 > t_1$ nach Reizende gesendet werden. (Eine modellhaft vereinfachende Beschreibung dieser sinnes- und nervenphysiologischer Vorgänge für informationspsychologische Zwecke wurde in Frank, 1959, S. 21 f., und, ausführlicher, 1962, S. 69 - 87, versucht.) Eine gewisse Reiznachwirkungszeit $t_s^* = t_2 - t_1$ besteht daher schon auf der Stufe der Sinnesorgane. Eine zusätzliche Trägheit kann vom anschließenden Neuronensystem einschließlich der Projektionszentren unterstellt werden. Apperzipierbar ist also nicht der Reiz selbst während der Zeit d, sondern seine Abbildung in einem zentralnervösen „Ultrakurzspeicher" in der um eine Gesamtnachwirkungszeit t_s verlängerten Zeit

$$(2)\ t = d + t_s \geq d + t_s^* > d$$

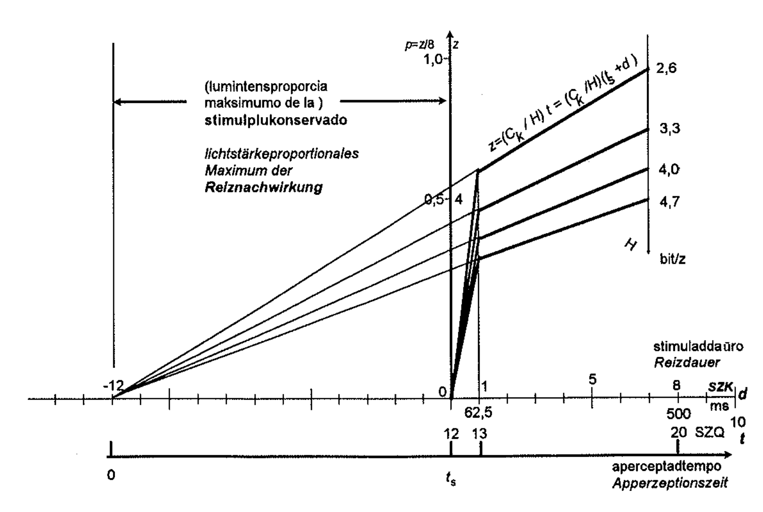


Bild 4: Theoretische Deutung von Bild 3A. Zur Reizdauer d kommt eine von der Redundanz, also von H, unabhängige Nachwirkungszeit t_s der Netzhaut und des anschließenden Perzeptionskanals (bis zu den sensorischen Projektionszentren bzw. zum Akkomodator) für die Apperzeption hinzu. Die apperzipierte Zeichenzahl z ergibt sich als Quotient aus apperzipierter Information $C_K t$ durch die (mittlere) Information H eines Zeichens.

Durch das wahrnehmungspsychologische Experiment von Averbach und Coriell (1961) wird dies bestätigt. Diese Experimentatoren projizierten *unmittelbar nach* einem Bild, das für die auch nur näherungsweise vollständige Apperzeption viel zu informationsreich war, an unvorhergesehener Stelle des (inzwischen verschwundenen) Bildes ein kleines Quadrat; ihre Versuchspersonen *apperzipierten dann noch nachträglich* alles, was sie zuvor an der so umrahmten Stelle *perzipiert* hatten.

Die in Abschnitt 2 aufgeführte Bedingung (c) für die psychologische Relevanz des Informationsmaßes ist also trotz des in Bild 3 dargestellten empirischen Befunds als erfüllt annehmbar, wenn als Wahrnehmungszeit statt der Reizdauer d die nach (2) verfügbare

Apperzeptionszeit t eingesetzt wird (Bild 4). In diesem Koordinatensystem gilt dann - statt der (in Bild 4 fett gezeichneten) empirisch beobachtbaren, geknickten Funktion $I(d)$ - die Proportionalitätsbeziehung

$$(3)\ I = C_K t = C_K(d+t_s)$$

Dabei ist C_K die zu bestimmende, als Apperzeptionsschnelle voll genutzte Kanalkapazität C vom Akkomodator zum Kurzspeicher. (Für den Fall des M/B/P-Versuchs wurde t_s zu rund ¾ Sekunden geschätzt.)

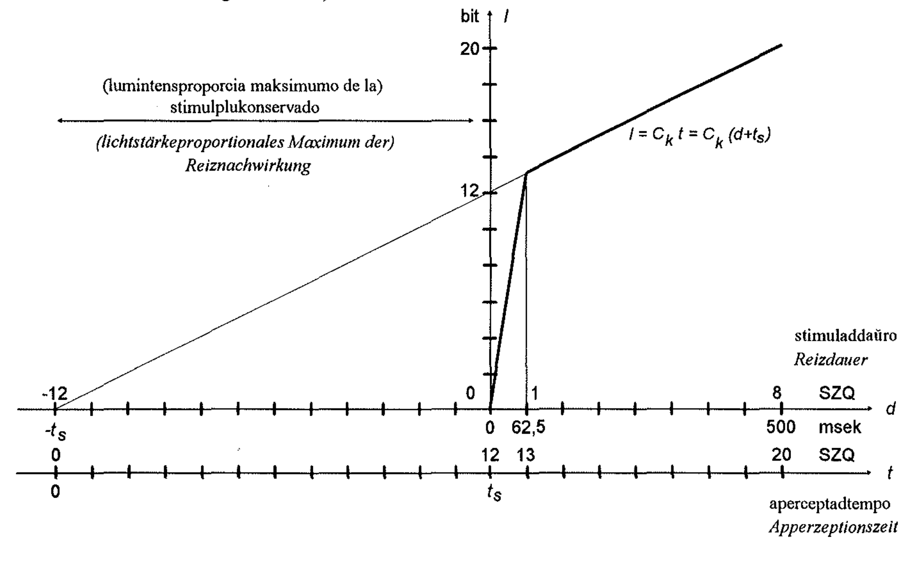


Bild 5: Idealisierung von Bild 3B. Die Reiznachwirkungszeit, *nicht aber* C_K, steigt mit der Reizintensität. Bei unterschiedlicher Reizintensität (z. B. bei unterschiedlicher Entfernung der Versuchspersonen von der Projektionsfläche, wie bei der Nachuntersuchung von Frank und Wagner, 1982) wird also der zweite Ast der M/B/P-Funktion *parallel* verschoben, während beim ersten Ast die Steigung sich ändert (also die Perzeptionsgeschwindigkeit).

Zur Erklärung des empirisch beobachtbaren Kurvenverlauf $I(d)$ für $d < 1$ SZQ müßte in diesem Bereich die verfügbare Apperzeptionszeit mit der Dauer d schneller wachsen als durch (2) erklärt wird; also dürfte t_s keine Konstante sein, sondern eine mit d monoton wachsende Funktion $t_s(d)$; die in Bild 4 hypothetisch gezeichnete Ursprungsgerade gilt im Spezialfall der Proportionalität:

$$(4)\ t_s(d) = \tau \cdot t_s(1\ \text{SZQ}) =: \tau \cdot t_s\ ,\quad d \leq 1\ \text{SZQ},\ 0 \leq \tau = d/\text{SZQ} \leq 1$$
$$(5)\ t = d + t_s(d) = d + t_s d/SZQ = d(1+t_s/\text{SZQ})$$

Bild 5 macht augenfällig, wie $t_s(d)$ als monoton steigende Funktion von d unter der Voraussetzung der Gültigkeit von (3) allgemein aus $I(d)$ zu ermitteln ist, und daß (nur) speziell im Falle der Proportionalität zwischen I und $d < 1$ SZQ hier auch die Nachwirkungszeit zu d proportional ist.

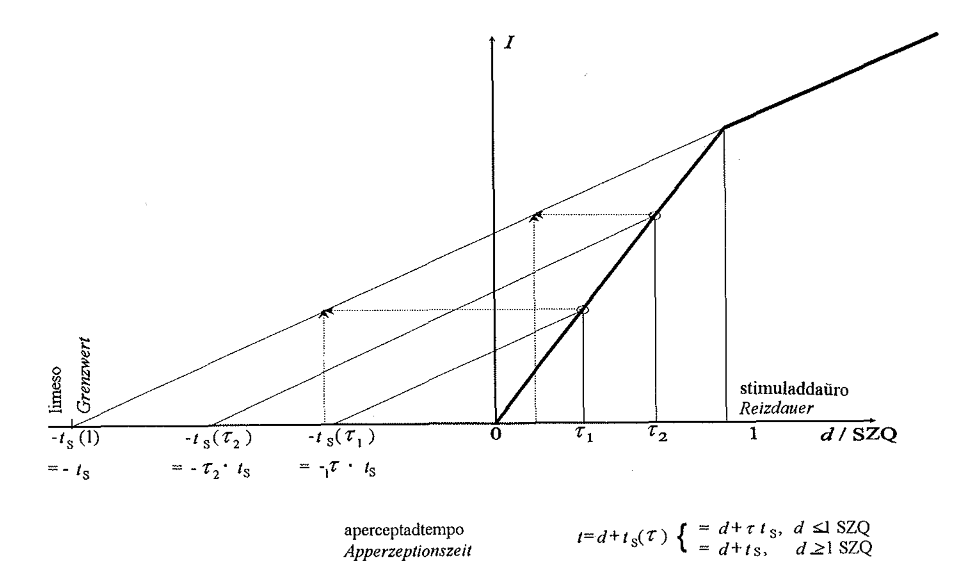


Bild 6: Im Falle *konstanter* Reizstärke wächst (beobachtbar) die apperzipierte Information, also die (theoretisch erschließbare) Apperzeptionszeit t proportional zur Reizdauer $d < 1$ SZQ. Sie setzt sich additiv aus dieser Reizdauer und der zu ihr bis $d = 1$ SZQ proportionalen Nachwirkungszeit zusammen. Der Knick der Funktion $I(d)$ entfällt daher bei der Funktion $I(T)$: die apperzipierte Information ist proportional zur Apperzeptionszeit.

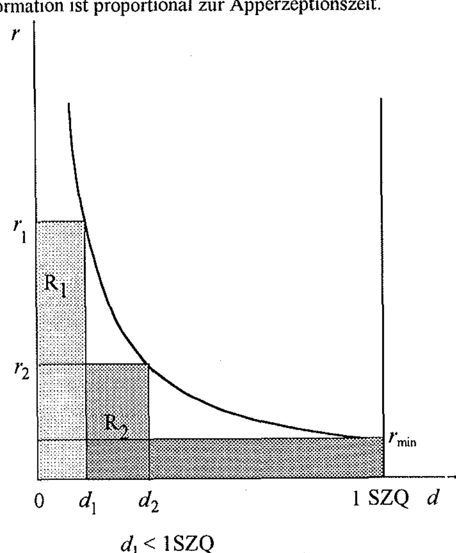


Bild 7: Sinnesphysiologische Erklärung von Bild 6. Bei gleichem Produkt aus Reizintensität und Reizdauer (< 1 SZQ) werden Reize nicht als unterschiedlich apperzipiert. Doppelte Reizdauer ist also zu doppelter Reizintensität äquivalent und führt zu längerer (doppelter?) Reiznachwirkungszeit.

Schon vorwissenschaftlich ist zu erwarten, daß der Grenzwert t_s(1 SZQ) der Nachwirkungszeit für $d \rightarrow 1$ SZQ auch für dasselbe Sinnesorgan und dieselbe Versuchsperson keine Konstante ist, sondern mit der Reizstärke steigt - und damit nach Konstruktion (Bild 5) auch jede kürzere Nachwirkungszeit $t_s(d)$, $d < 1$ SZQ: man „sieht" nach Schließen der Augen eine Lichtquelle desto länger, je heller sie ist. Dies wird durch den wahrnehmungspsychologischen Sachverhalt bestätigt, daß von zwei überschwellig starken, weniger als 1 SZQ dauernden Reizen nicht sicher empfunden werden kann, welches der längere bzw. der stärkere ist (Bild 6). *Subjektiv* vergleichbar sind vielmehr nur die Produkte $R = d \cdot r$ dieser beiden Reizdimensionen. Beim M/B/P-Versuch kann unterstellt werden, daß bei den Teilversuchen mit Buchstabenketten *unterschiedlicher Informacio* bzw. Redundanz die *Lichtstärke unverändert* blieb. Da Redundanz und Informacio auf der Stufe der Perzeption noch nicht relevant sein können, muß daher bei gleicher Expositionszeit d die Nachwirkungszeit t_s für alle vier Teilversuche dieselbe sein - und zwar sowohl im Grenzwert (Bild 7), als auch (nach der in Bild 6 dargestellten Konstruktion) für $d < 1$ SZQ. Dann aber ist nach (2) die Apperzeptionszeit nur von d, nicht auch von der Redundanz bzw. der Informacio der während der Zeit d gezeigten Buchstaben-8-tupel abhängig. Die Übereinstimmung der in Informacio-Einheiten (bit) gemessenen Information, die in dieser Zeit apperzipiert wird, besagt dann, daß auch für Expositionszeiten $d < 1$ SZQ, also für Apperzeptionszeiten $t < t_s + 1$ SZQ, die Relevanzbedingung (b) aus Abschnitt 2 erfüllt ist, also (3) nicht erst für größere Apperzeptionszeiten gilt. (Etwas anderes wäre

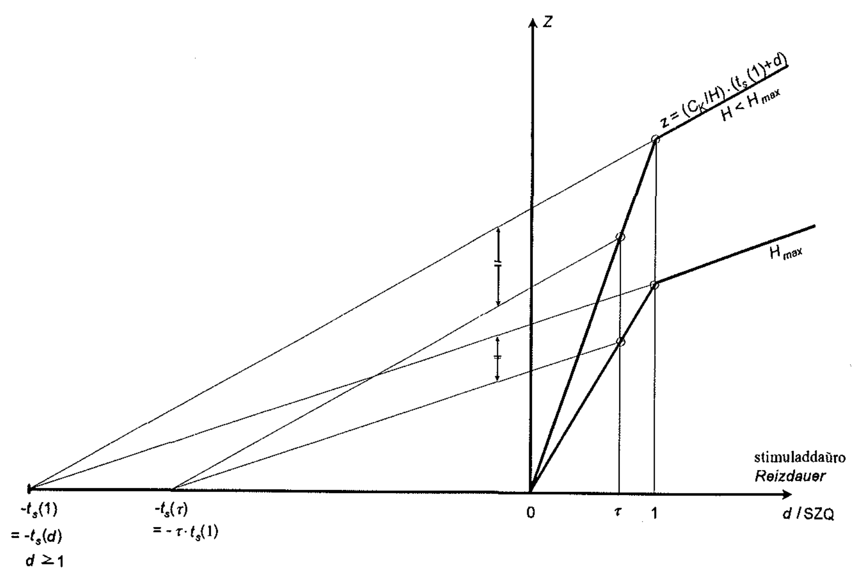


Bild 8: Bei gleicher Reizstärke hängt die Reiznachwirkungszeit nur von der Reizdauer $d < 1$ SZQ bzw. vom Bruchteil $\tau = d$/SZQ ab, nicht vom Informationsgehalt der Zeichen. Soweit die Reiznachwirkungszeit proportional zur Reizdauer ist, verlaufen auch die ersten Äste der M/B/P-Funktion (Bild 3A) geradlinig, und ihre Steigung ist zu H umgekehrt proportional. *Genau dies zeigen die Versuchsergebnisse von M/B/P besonders eindrucksvoll, ohne daß es bisher theoretisch aufgearbeitet worden wäre.*

kaum zu erklären, wenn man die Apperzeption [Bewußtwerdung] als Übertragung durch den Kanal vom Akkomodator zum Kurzspeicher versteht.) Damit ist das Zusammenfallen der empirischen Kurven auch für $d < 1$ SZQ in einer mit der Beobachtung für $d > 1$ SZQ übereinstimmenden Weise erklärt, und die empirischen Ergebnisse dieser Versuche mit sehr kurzen Reizen bestätigen die Kohärenz des informationspsychologischen Ansatzes.

5. *Theoriegeleitete Auswertung der M/B/P-Daten.*

Ohne zunächst den „informationspsychologischen Fundamentalsatz" (1) und die Nachwirkungsannahmen (4; 5) schon zu unterstellen, berechneten wir aus den M/B/P-Originaldaten die Regressionslinien für die Meßwerte in den Projektionszeitintervallen bis 40 msek bzw. ab 100 msek. Sie schneiden sich an der Stelle $d_P = 45{,}3$ msek und schneiden die Zeitachse bei $d = -1{,}8$ msek bzw. $d = -1068$ msek ($= -t_S$). Der Ordinatenwert des Schnittpunkts (in bit) und die Steigungen der beiden Regressionslinien (in bit/sek) hängen vom Informationsgehalt $8H_{max}$ der redundanzfreien Zeichen-8-tupel ab. Rechnet man H_{max} /bit = ld 27 ≈ 4,75, dann erhält man als Ordinatenwert des Knickpunkts $I_P = 14{,}3$ bit und als Steigung der ersten Regressionslinie (Perzeptionsschnelle) ca. 304 bit/sek, für den zweiten Ast $C_K \approx 12{,}85$ bit/sek. Unterstellt man jedoch eine Binärcodierung der 27 Buchstaben durch je 5 Bit oder - sparsamst - von 22 Buchstaben durch je 5 Bit und der restlichen 5 durch je 4 Bit, also im Mittel durch ca. 4,81 Bit, dann erhöhen sich Ordinaten- und Steigungswerte um 5,15% bzw. 1,26%. Mit der letzteren Annahme, die am besten zum Modellansatz paßt, erhöht sich C_K auf 13,01 bit/sek. Die Annahme

$$(6)\ I(d) = C \cdot d \quad \text{für } d < 1 \text{ SZQ}$$

ist also mit den Daten sehr gut verträglich. Die Steigung des ersten Astes kann aus der als gesichert geltenden Annahme $I(0) = 0$ daher genauer mit

$$(7)\ C = \Sigma\, I_i d_i \,/\, \Sigma\, d_i^2$$

zu etwa 322 bit/sek ermittelt werden. (Diese Gleichung entsteht aus der Formel für die Steigung der Regressionsgeraden, wenn man den theoretisch gesicherten Punkt [0; 0] n-fach den Meßpunkten zufügt und dann $n \rightarrow \infty$ gehen läßt.) Damit erhält man für den Knickpunkt den verbesserten Wert $d_P = 44{,}4$ msek.

Unterstellt man nun die Gültigkeit von (1), dann würde für die Versuchspersonen von M/B/P 1 SZQ = 1/13,01 sek = 76,8 msek > 44,4 msek = d_P gelten. Die unbefriedigende Abweichung dieser beiden Meßergebnisse für das SZQ kann durch die (im Vergleich zum ersten Ast) relativ große Streuung der Meßwerte des zweiten Asts der M/B/P-Funktion erklärt werden, die eine nur ungenaue Ermittlung der Steigung C_K ermöglicht. Unterstellt man eine etwas größere Steigung, dann sinken der Kehrwert $1/C_K$ sowie die Knickpunktstelle d_P , und man kann iterativ den Koinzidenzwert ermitteln. Dazu betrachten wir den ersten Ast der Kurve als gesichert. Bei der Regressionsrechnung für den zweiten Ast ist damit der Knickpunkt (d_P; I_P) als theoretisch gesicherter Punkt hinzuzufügen. C_K ist nach Koordinatenverschiebung durch diesen Knickpunkt nach (7) berechenbar, wenn dort I_i

durch I_i - I_P und d_i durch d_i - d_P ersetzt wird. Kriterium dafür, daß der mit den Daten bestverträgliche Wert d_P in die Regressionsrechnung eingegeben wurde, ist, daß dieses d_P mit dem Kehrwert $1/C_K$ der so berechneten Steigung der Regressionslinie numerisch zusammenfällt.

Für 1 SZQ = 0,0400 sek erhält man bei dieser Rechnung C_K = 24,9 bit/sek, also einen Apperzeptionszeitbedarf von 0,0401 sek/bit. Die (studentischen) Versuchspersonen von M/B/P zeichneten sich demnach durch eine erheblich über dem Durchschnitt liegende Apperzeptionsschnelle und ein entsprechend feines Zeitauflösungsvermögen aus. Dem entspräche nach Lehrl (1981) ein IQ von ca. 140.

Schrifttum

Averbach, E., & A. S. Coriell: *Short-Term Memory in Vision.* The Bell System Technical Journal, 1961, 309 - 328.

Barandovská, Vera: *Kybernetische Pädagogik / Klerigkibernetiko.* Band 6 (Schriften 1973 - 1992), Band 7 (Schriften 1962 - 1992). AL durch Esprima Bratislava, Kava-Pech Prag, und IfK Berlin & Paderborn, 1993.

Cherry, Colin : *On Human Communication.* The Technology Press of Massachusetts Institute of Technology & J. Wiley, New York, 1957.

Cisteian Eugenia Maria: *Noto pri la progresrapideco de la informacipsikologio.* GrKG/Humankybernetik 36/4, 1995, 151 - 154.

Faloba Diana: *Aliĝo al pli profunda teoria ekspluato de la empiriaj rezultoj de Miller, Bruner kaj Postman el 1954.* Bakkalaureatsarbeit, vorgelegt der Sektion Kybernetik der AIS, 1997.

Frank, Helmar: *Informationsästhetik. Grundlagenprobleme und erste Anwendung auf die mime pure.* Dissertation TH Stuttgart, 1959 (Verlag: Heß, Waiblingen). Schnelle, Quickborn, [2]1968. (Nachdruck in Meder / Schmid, 1974, 1 - 114.)

Frank, Helmar: *Über eine informationspsychologische Maßbestimmung der semantischen und pragmatischen Information.* GrKG /H. 1 /2, 1960, 37 - 40.

Frank, Helmar: *Zum Problem des vorbewußten Gedächtnisses.* GrKG/H. 2/1, 1961, 17-24. (Nachdruck in Meder / Schmid, 1973, Bd. 1, 35 - 44.)

Frank, Helmar: *Kybernetische Grundlagen der Pädagogik. Eine Einführung in die Informationspsychologie und ihre philosophischen, mathematischen und physiologischen Grundlagen.* Agis, Baden-Baden, & Gauthier-Villars, Paris, 1962. (Nachdruck in Barandovska, 1993, Bd. 7, 1 - 188.) [2]1969, 2 Bände: Agis, Baden-Baden, & Kohlhammer, Stuttgart. [3]1971, gekürzte Taschenbuchausgabe (Hrsg.: B. S. Meder): Kohlhammner, Stuttgart.

Frank, Helmar: *Informationspsychologie und Nachrichtentechnik.* In: Norbert Wiener & J.P. Schadé (Hsg.): *Nerve, Brain and Memory Models.* Elsevier Publishing Comp., Amsterdam, 1963, 79 - 96. (Nachdruck in Meder / Schmid, 1973, Bd. 2, 101 - 118.

Frank, Helmar: *Das Bewußtsein als „Ding an sich" der Kybernetik.* In: H.-W. Klement (Hsg.): *Bewusstsein. Ein Zentralproblem der Wissenschaften.* Agis, Baden-Baden, 1975, 51 - 61; 251 - 252.

Frank, Helmar: *Denkanstöße zum gemeinsamen Selbstverständnis und zur gemeinsamen Selbstdarstellung der Kommunikationskybernetiker* . In: M.Krause & S. Piotrowski (Hsg.): *Bildungskybernetik und Europäische Kommunikation.* Akademia Libroservo durch KAVA-PECH, Prag, 1993, 191 - 195.

Frank, Helmar: *Plädoyer für eine Zuziehung der Semiotik zur Kybernetik.* GrKG/H. 36 / 2, 1995, 61 - 72.

Frank, Helmar: *Kommunikationskybernetik - das theoretische Fundament der Bildungstechnologie.* In Siegfried Piotrowski (Hsg.): *Kybernetische Ursprünge der europäischen Bildungstechnologie.* Akademia Libroservo durch IfK-Verlag, Paderborn, 1996(a), 40 - 52.

Frank, Helmar: *Klerigkibernetiko / Bildungskybernetik* . Akademia Libroservo durch ESPRIMA Bratislava & KoPäd München, 1996(b).

Frank, Helmar, & Herbert W. Franke: *Ästhetische Information / Estetika Informacio.* AL durch IfK-Verlag Paderborn & KoPäd München, 1997.

Frank, H., H. Jeske & S. Lehrl: *„Faustformeln" zum IQ* . GrKG/Humankybernetik 23/1, 1982, 23 - 28. (Nachdruck in Barandovská, 1993, Bd. 7, 577 - 582.)

Frank, Helmar, & Brigitte S. Meder: *Einführung in die kybernetische Pädagogik.* Deutscher Taschenbuch-Verlag , München ,1974. (Nachdruck in Meder / Schmid, 1974, 381 - 584.)

Frank, H., & H. Wagner: *Messung der Apperzeptionsgeschwindigkeit mit einem Experimentalfilm.* GrKG/H. 23/2, 1982, 73 - 79. (Nachdruck in Barandovská, 1993, Bd. 6, 609 - 615.)

Hilgers, Rainer, & Yashovardhan (Hrsg.): *EG-Wörterbuch mathematischer Begriffe / EK-Vortaro de matematikaj terminoj.* Leuchtturm, Alsbach, 1980.

Lehrl, Siegfried: *Hatte Francis Galton doch recht?* GrKG/H. 22/1, 1981, 17 - 28.

Meder, Brigitte S., & Wolfgang F. Schmid (Hsg.): *Kybernetische Pädagogik. Schriften 1958 - 1972.* Kohlhammer, Stuttgart, & IfK, Paderborn. Bände 1 - 4, 1973; Bd. 5, 1974.

Miller, George A, Jerome S. Bruner & Leo Postman: *Familiarity of letter sequences and tachistoscopic identification..* The Journal of General Psychology, 50, 1954, 129 - 134.

Rein, Hermann, & Max Schneider: *Einführung in die Physiologie des Menschen.* Springer, Heidelberg, [11]1955 .

Riedel, Harald : *Empirische Untersuchungen zur Kybernetischen Pädagogik.* Schnelle, Quickborn, 1965.

Roth, Erwin , W. D. Ostwald & K. Daumenlang: *Intelligenz* .W.Kohlhamer, Stuttgart, 1972; [4]1980.

Schmidt, Hermann: *Denkschrift zur Gründung eines Institutes für Regelungstechnik.* VDI-Druck, Berlin, 1941. (Nachdruck als Beiheft der GrKG/H. 2/1961.)

Shannon, Claude E.: *A mathematical theory of communication.* Bell Syst. Techn. J. 27, 1948, 379 - 423; 623 - 656. (Nachdruck in C. E. Shannon & W. Weaver: *The Mathematical Theory of Communication.* The University of Illinois Press: Urbana 1949, 3 - 91.)

von Buddenbrock, W.: *Vergleichende Physiologie.* Band 1. Birkhäuser, Bassel , 1952.

von Buddenbrock, W.: *Die Welt der Sinne. Eine gemeinverständliche Einführung in die Sinnesphysiologie.* Springer, Heidelberg, [2]1953.

von Cube Felix: *Kybernetische Grundlagen des Lernens und Lehrens.* Klett, Stuttgart, 1965.

Weltner Klaus: *Informationstheorie und Erziehungswissenschaft.* Schnelle, Quickborn, 1970.

Eingegangen am 4. August 1997

Anschriften der Verfasser: Prof.Dr.Helmar Frank, Kleinenberger Weg 16A, D-33100 Paderborn. Diana Faloba, Bahluiului 7A , RO-2400 Hermannstadt/Sibiu

Provo kompletige ekspluati la ŝlosileksperimenton de la informacipsikologio. (Resumo).

Miller, Bruner kaj Postman montris 1954, ke dum certa daŭro de taĥistoskopa prezentado oni ja malpli da *procentoj de litersinsekvo* enprenas (kaj povas reprodukti), se la litersinsekvo enhavas pli da informacio, sed ke la tiel ricevebla *informacio* estas la sama (bildoj 1 - 2). El la mezurvaloroj por prezentadtempoj $d \geq 1$ subjektiva tempero (SZQ) la informacipsikologio konkludis sian „fundamentan teoremon", laŭ kiu la aperceptadrapideco estas 1 bit/SZQ. La kunfandiĝo de la empiriaj kurboj post transkalkulo de signoprocentoj (bildo 1) al informacivaloroj (bildo 2) estas pli okulfrapa por $d < 1$ SZQ. Pro la konata, supozeble proporcia kresko de la stimulkonservadtempo sur la retino kaj en la sekva perceptadkanalo, la aperceptadtempo superas je la stimulkonservadtempo la stimuladtempon kaj ebligas apercepti proporcian kvanton da informacio jam dekomence. Ĉi tiuj teoriaj konsideroj ebligas pli fajne ekspluati la empiriajn datumojn de Miller, Bruner kaj Postman. Montriĝis, ke siatempe estis je dispono esploritoj kun alta aperceptadrapideco (preskaŭ 24 bit/sek), do - laŭ Lehrl - kun alta inteligenteckvociento (preskaŭ 140).

grkg / Humankybernetik
Band 38 · Heft 3 (1997)
Akademia Libroservo / IfK

Didaktischer Test vom statistischen Gesichtspunkt betrachtet

Stanislav KOMENDA, Olomouc (CZ)

aus der Medizinischen Fakultät der Palacký Universität Olomouc, Tschechien

1. *Problemstellung. Die Rückkoppelung*

Das Lernen spielt eine so wichtige Rolle im menschlichen Leben, daß 15 bis 25 Prozent dieser mittleren Lebenslänge mit Lernen erfüllt und fast alles weitere Leben vom Lernen durchdrungen ist. Eine notwendige Komponente des Lernens stellt die Rückkoppelung dar, die den Informationsfluß über die Kenntnisse des lernenden Subjekts sicherstellt. Solche Information soll durch die Beobachtung und das Messen gewonnen werden. Die genaue Darstellung der Kenntnisse der Schüler stellt eine notwendige Bedingung für die Leistungsfähigkeit eigener Tätigkeit des Lehrers dar, denn nur dann kann er diese Tätigkeit flexibel modifizieren und anpassen. Die Prüfung aller Art, Testen der Kenntnisse ist praktisches, alltägliches Messen im Dienste der Rückkoppelung und Regelung eines pädagogischen Prozesses.

2. *Induktive Auffassung*

Die Kenntnisse sind nicht erschöpfend zu messen; dazu genügt es zu bemerken, daß auch das Testen allein das Gemessene beeinflußt. Die Kenntnisse zu messen - es bedeutet nur eine Auswahl aus dem möglichen Raum dem Prozeß des Testen zu unterwerfen - was notwendigerweise von Bemühungen die partiellen Befunde für eine breitere Situation zu verallgemeinern begleitet ist. Aus den Antworten auf eine kleine Fragenmenge sind die Schlußfolgerungen über die komplexen Kenntnisse des Ganzen zu ziehen. Prüfen - bedeutet generalisieren, verallgemeinern. Prüfen bedeutet Induktion - mit allen Problemen, die einen solchen Prozeß notwendigerweise begleiten: nämlich den Irrtum zu riskieren, daß die existierenden Kenntnisse nicht erkannt werden oder die Ignoranz nicht diagnostiziert wird. Es gibt eine wissenschaftliche Disziplin, die sich ausschließlich mit den Problemen der Induktion beschäftigt, die *mathematische Statistik.* Dieser Gegenstand existiert schon seit hundert Jahren, mit Anwendung von statistischen Methoden in der Psychologie, Biologie, Medizin wie auch in der Ökonomie - und es gibt keine Argumente, warum sie auch auf dem Gebiet der Pädagogik nicht angewendet werden dürften.

Wir sollten aber etwas präziser sein - die Methoden der mathematischen Statistik sind in der pädagogischen Domäne nicht unbekannt. Eher handelt es sich um die Häufigkeit der Anwendungen und um eine Ansiedlung der statistischen Methoden in dem Instrumen-

tarium der pädagogischen Forschung. Es handelt sich also darum, daß die statistischen Konzepte zum Ausgangspunkt für die Auffassung der pädagogischen Phänomene und zu Bestandteilen der Struktur der pädagogischen Begriffe werden. Es geht darum, daß - analogisch zur Psychometrie, Anthropometrie, Soziometrie, Ökonometrie und Biometrie - auch der Inhalt des Gegenstandes, der mit dem Namen Edukometrie bezeichnet werden sollte, erfüllt werde.

3. *Die Kenntnisse und deren Prüfung*

In unserem Konzept sind drei Stufen der Kenntnisse zu unterscheiden:

(a) Der Raum der wirklichen, aktuellen Kenntnisse des Subjekts von den zu prüfenden Themen. Die Tiefe der Kenntnisse wird im Sinne der Bloomschen Taxonomie verstanden. Diese Stufe ist nicht durch eine direkte Messung ergreifbar - und darf als Anteil des nicht beherrschten Lehrstoffes konzipiert werden.

(b) Die Stufe der wirklichen Kenntnisse, die das Subjekt von dem geprüften Thema hat, äußert sich indirekt durch die Lösungen der Testprobleme (Testfragen), die eine (zufällige) Auswahl aus der Menge aller möglichen Probleme darstellen. Es ist deswegen sinnvoll von der Ebene der Kenntnisse, die das Subjekt von den Testfragen hat, zu sprechen. Wenn sich die Prüfung auf einem multiple-choice Test gründet, manifestiert sich diese Ebene durch einen Wert an der Testskala (Score der falschen Antworten). Der entsprechende Wert ist die einzige für den Lehrer zugängliche Informationsquelle von den Kenntnissen des Subjekts. Die Effizienz dieser Quelle wächst mit der Verbundenheit der Antwort mit den wirklichen Kenntnissen des Subjekts. Diese Verbundenheit ist statistisch ergreifbar - in der Form der entsprechenden Wahrscheinlichkeitsverteilungen - und deswegen sollten auch diese Kenntnisse statistisch gemessen werden.

Die bedingten Wahrscheinlichkeitsverteilungen sind als die Informationskanäle zu begreifen, durch die die Information aus dem Raum der wirklichen Kenntnisse zum Raum der ergreifbaren Antworten des Subjekts hinfließt.

(c) Der zufällige Mechanismus der Auswahl der Testfragen im Raum der wirklichen Kenntnisse wird auch damit modifiziert, daß wegen der Auswahl der Antworten aus der für jede Testfrage angebotenen Menge der Antwortalternativen - auch das Subjekt, das von dem Teilproblem nichts weiß, die richtige Antwort mit einer positiven Wahrscheinlichkeit erreichen kann. Die Intensität dieses Faktors hängt davon ab, wieviel unrichtige Alternativen (Distraktoren) zur Auswahl des Subjekts mit der Frage angeboten wurden [2,5].

4. *Parameter des Tests*

Hiernach soll die Situation eines Subjekts im Test untersucht werden mit dem Ziel, den Anteil der bei ihm wirklich vorhandenen (bzw. der nicht vorhandenen) Kenntnisse zu beurteilen. Der Test möge n Fragen (Problemen, Aufgaben) enthalten, die als voneinander unabhängig angenommen werden. Wir setzen also voraus, daß die Kenntnis der richtigen Lösung auf jede Aufforderung keinen Einfluß auf die Möglichkeit habe, auf die jeweils übrigen Aufforderungen des Tests richtig zu reagieren.

Durch den Test wird dem Subjekt zu jeder Aufforderung eine Menge von q Antworten (Alternativen) zur Auswahl angeboten. Der Test ist so konstruiert, daß nur eine einzige Antwort richtig ist; q-1 Antworten sind also immer falsch (und in der Rolle der Distraktoren angewendet). Dieser Umstand sei dem Subjekt bekannt. Es führt dazu, daß die richtige Antwort auch zufällig, durch Raten, gefunden werden kann.

5. *Die Kenntnisse und das Testergebnis des Subjekts*

Wir bezeichnen mit dem Symbol π den Anteil des dem Subjekt unbekannten Lehrstoffes bezüglich einer Aufforderung ($0 \leq \pi \leq 1$) siehe Abb.1.

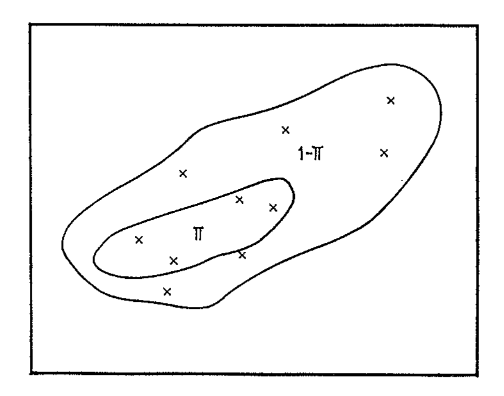


Abb. 1

Dieser Parameter soll mit Hilfe des Testergebnisses geschätzt werden. Das Symbol p bedeutet die Wahrscheinlichkeit, daß das Subjekt auf eine zufällig ausgewählte Aufforderung eine falsche Antwort gibt. Nach dem Satz für die totale Wahrscheinlichkeit erhält man als Beziehung zwischen den Größen π und p [2,4]

$$1-p=(1-\pi)\cdot 1+\pi\frac{1}{q} \quad \text{bzw.}$$

$$(1) \qquad p=\pi\frac{q-1}{q}$$

Daraus folgt, daß - wegen der Möglichkeit eines zufälligen Erfolges des ignoranten Subjekts - die Wahrscheinlichkeit p einer falschen Antwort kleiner ist als der Anteil π des unbekannten Lehrstoffes.

Aus (1) läßt sich die inverse Wahrscheinlichkeit dafür berechnen, daß das Subjekt, das auf die zufällig ausgewählte Testfrage richtig geantwortet hat, auch wirklich die entsprechenden Kenntnisse hatte. Diese Wahrscheinlichkeit ist (dem Satz nach Bayes gemäß)

$$P(K|R)=\frac{P(R|K)\,P(K)}{P(R|K)\,P(K)+P(R|\overline{K})\,P(\overline{K})}$$

wobei die Symbole folgende Ereignisse bezeichnen

K ... das Subjekt kennt das Problem

$\overline{K}$ ...das Subjekt kennt das Problem nicht

R ... das Subjekt antwortet richtig

so daß

$$P(K|R)=\frac{1\cdot(1-\pi)}{1\cdot(1-\pi)+\frac{1}{q}\pi}$$

$$=\frac{1-\pi}{1-\pi+\frac{1}{q}\pi}=\frac{q(1-\pi)}{\pi+q(1-\pi)}.$$

Daraus ergibt sich die komplementäre Wahrscheinlichkeit dafür, daß bei richtiger Beantwortung einer Frage keine Stoffkenntnis vorhanden ist (die richtige Antwort nur zufällig erraten wurde)

$$\frac{\pi}{\pi+q(1-\pi)}.$$

Durch diese Methode der inversen Wahrscheinlichkeit kann man auch feststellen, daß ein Subjekt, das eine zufällig ausgewählte Frage falsch beantwortet hat, mit der Wahrscheinlichkeit Eins das entsprechende Wissen nicht besitzt.

6. *Die Kenntnisse und die Anzahl der unbekannten Fragen*

Unter der Bedingung, daß der Anteil π des Stoffes nicht bekannt ist, läßt sich die Wahrscheinlichkeit dafür, daß l Aufforderungen ($0 \leq l \leq n$) des Testes in diesen unbekannten Bereich fallen, ermitteln mit

$$(2) \qquad P(l\,|\,\pi,n)=\binom{n}{l}\pi^{l}(1-\pi)^{n-l}$$

Diese Beziehung wurde unter der Bedingung einer gleichmäßigen Verteilung der Aufforderungen (Testfragen) über das geprüfte Stoffgebiet abgeleitet. Die Situation ist in Abb.1 für den Fall n=10 und l=4 dargestellt.

7. *Die Anzahl der unbekannten und falsch beantworteten Fragen*

Die Aufforderungen aus dem Bereich des vorhandenen Wissens werden mit Sicherheit richtig beantwortet. Daneben kann es einige Fragen aus dem Bereich des nicht vorhandenen Wissens geben, die ebenfalls richtig beantwortet werden. Die Zusammenhänge lassen sich schematisch folgendermaßen darstellen (siehe Tabelle 1). Die Wahrscheinlichkeit, daß unter den l Fragen aus dem Bereich des nicht vorhandenen Wissens k Fragen falsch und l-k zufällig richtig beantwortet werden, läßt sich durch die Binomialwahrscheinlichkeit

$$(3)\qquad P(k\,|\,l,q)=\binom{l}{k}\left(\frac{q-1}{q}\right)^{k}\left(\frac{1}{q}\right)^{l-k}$$

ermitteln. Dabei handelt es sich um die bedingte Verteilung der Zufallsveränderlichen X (Anzahl der Fehler im Test) unter der Bedingung, daß l Aufforderungen des Tests in den Bereich der vom Subjekt nicht beherrschten Kenntnisse gefallen sind, X=0,1,...,l.

Die Zufallsvariablen X (die Fehleranzahl) und Y (die Anzahl der Fragen aus dem Bereich des nicht vorhandenen Wissens) unterscheiden sich wesentlich: die X-Werte sind

k falsche Antworten		n-k richtige Antworten
l Fragen aus dem Bereich des nicht vorhandenen Wissens	n-l Fragen aus dem Bereich des vorhandenen Wissens	
n Aufforderungen des Tests		

Tabelle 1

direkt beobachtbar; die Y-Werte bieten keine Möglichkeit einer direkten Beurteilung. Für die theoretische Wahrscheinlichkeitsverteilung der Zufallsvariablen X erhält man

$$(4)\qquad \begin{aligned} P(X=k|\pi,n,q) &= \sum_{l=k}^{n} P(k|l,q)\; P(l|\pi,n) \\ &= \binom{n}{k}\left(\pi\frac{q-1}{q}\right)^{k}\left(1-\pi\frac{q-1}{q}\right)^{n-k} \end{aligned}$$

Wie man sieht, nähert sich die Wahrscheinlichkeitsverteilung (4) mit wachsendem q der Binomialverteilung (2). Der Zusammenhang dieser Variablen ist durch Abb. 2 dargestellt.

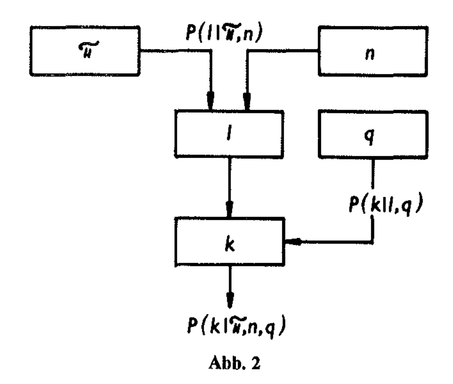


Abb. 2

8. *Die Maximum-Likelihood Schätzung des Parameters* π

Man kann beweisen, daß für π keine erwartungstreue Schätzung existiert [4]. Deshalb suchen wir nachfolgend eine Maximum-Likelihood Schätzung $\hat{\pi}$ von π in Abhängigkeit

von der Zahl k der falschen Antworten. Zur Ableitung wird der Wert $\hat{\pi}$ gesucht, für den (4) ein Maximum (im Intervall (0,1)) annimmt. Als Lösung erhält man (siehe Abb. 3)

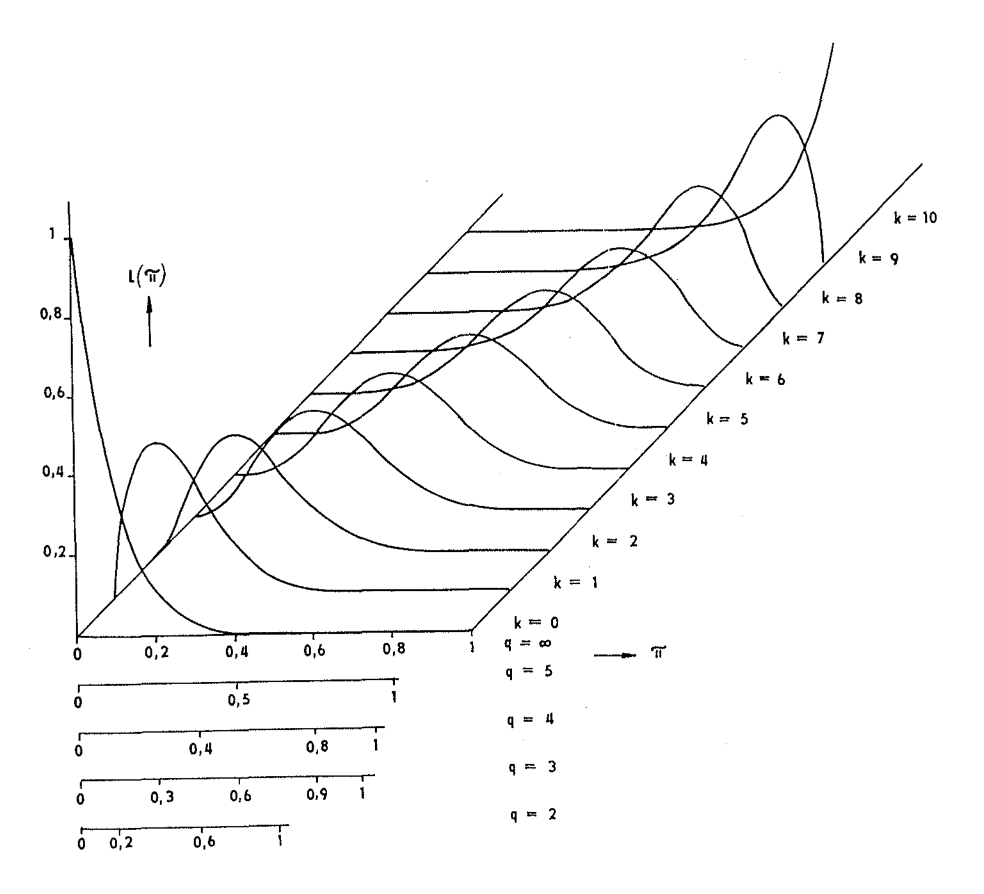


Abb. 3

$$(5)\qquad \begin{aligned} \hat{\pi} &= \frac{k}{n}\,\frac{q}{q-1}, &&\text{wenn } k \le n\,\frac{q-1}{q} \\ \hat{\pi} &= 1, &&\text{wenn } k > n\,\frac{q-1}{q} \end{aligned}$$

Abb. 4 zeigt den Verlauf der Schätzung für den Fall n = 20, q=4.

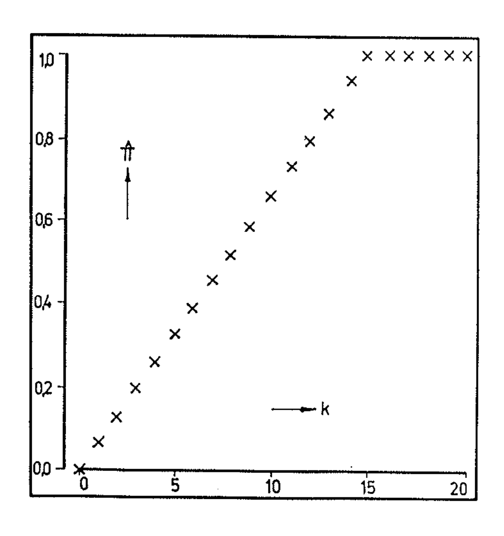


Abb. 4

9. *Der Mittelwert der Maximum-Likelihood Schätzung*
Für den Mittelwert der Schätzung (5) bekommen wir den Ausdruck

$$E(\hat{\pi}|\pi) = \pi B\left(\left[n\frac{q-1}{q}\right]-1 \mid n-1, \pi\frac{q-1}{q}\right) + \qquad (6)$$
$$+1 - B\left(\left[n\frac{q-1}{q}\right] \mid n, \pi\frac{q-1}{q}\right)$$

wobei $[x]$ den ganzen Teil von x und $B(k \mid n, p)$ die Distributionsfunktion der Binomialverteilung bedeuten. Die Schätzung (5) ist verzerrt (biased); der Mittelwert (6) fällt niedriger als der wahre Parameterwert π aus. Die Größe der Verzerrung konvergiert sowohl mit wachsendem Antwortsangebot q als auch mit zunehmendem Testumfang n gegen Null. Für den an der Abb. 4 gegebenen Fall ist der Verlauf des Mittelwertes in Abhängigkeit von π in Abb. 5 dargestellt.

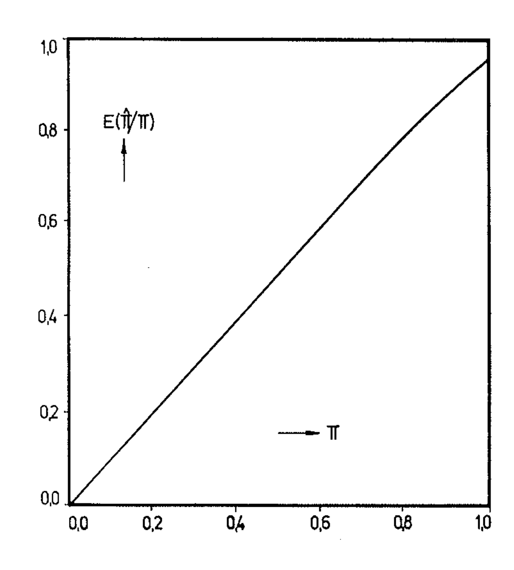


Abb. 5

10. *Die Streuung der Maximum-Likelihood Schätzung*
Die Streuung der Schätzung (5) berechnen wir mittels der allgemeinen Formel

$$D(\hat{\pi} \mid \pi) = E(\hat{\pi}^2 \mid \pi) - E^2(\hat{\pi} \mid \pi).$$

Nach einigen Berechnungen erhält man

$$D(\hat{\pi}|\pi) = \frac{n-1}{n}\pi^2 B\left(\left[n\frac{q-1}{q}\right]-2 \mid n-2, \pi\frac{q-1}{q}\right)$$
$$+\frac{1}{n}\frac{q}{q-1}\pi B\left(\left[n\frac{q-1}{q}\right]-1 \mid n-1, \pi\frac{q-1}{q}\right)$$
$$+1 - B\left(\left[n\frac{q-1}{q}\right] \mid n, \pi\frac{q-1}{q}\right) \qquad (7)$$
$$-\left\{\pi B\left(\left[n\frac{q-1}{q}\right]-1 \mid n-1, \pi\frac{q-1}{q}\right)\right.$$
$$\left.+1 - B\left(\left[n\frac{q-1}{q}\right] \mid n, \pi\frac{q-1}{q}\right)\right\}^2$$

Im Spezialfall eines didaktischen Tests mit freier Antwortwahl (ohne Antwortvorgabe, d.h. mit $q \to \infty$) reduziert sich (7) zu

$$(8) \qquad D(\hat{\pi}\,|\,\pi) = \frac{1}{n}\,\pi\,(1-\pi),$$

zur Streuung der Binomialverteilung. Die Abhängigkeit der Streuung (7) vom wahren Parameterwert π und der Zahl der Antwortvorgaben q ist für den Fall eines Tests mit einem

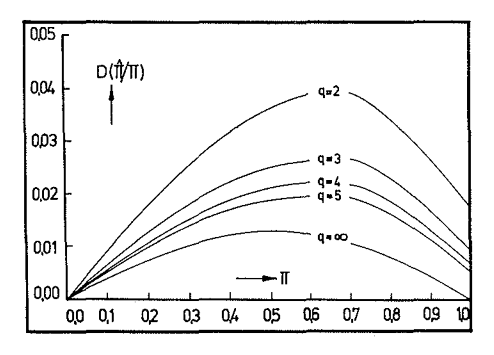


Abb. 6

Umfang von n = 20 in Abb. 6 dargestellt. Mit Hilfe der Streuung (7) kann man auch die Zuverlässigkeit eines didaktischen Tests bewerten: Die Genauigkeit der aus dem Testergebnis abgeleiteten Schätzung des wirklichen Niveaus der Kenntnisse eines Subjekts ist der Streuung dieser Schätzung umgekehrt proportional. Die Streuung sinkt mit zunehmendem Testumfang n und mit wachsender Anzahl der Antwortvorgaben. Daraus folgt, daß ein Test mit einer kleineren Anzahl von Antwortvorgaben im Vergleich zu einem Test mit größerem q einen größeren Testumfang n erfordert, um eine vorgegebene Genauigkeitsgrenze der Parameterschätzung zu erreichen. In Abb. 7 ist für den Fall, daß der Anteil des nicht beherrschten Stoffes am Stoffumfang des Tests n = 0,25 beträgt, die Abhängigkeit der Streuung von n und q veranschaulicht.

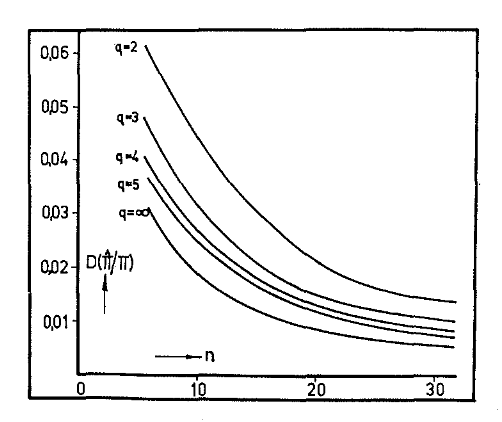


Abb. 7

10. *Bayessche Schätzung der Anzahl der auf den Bereich nicht beherrschten Stoffes entfallenen Fragen*

Die Zufallsveränderliche Y wurde als Anzahl der Testfagen eingeführt, die auf den Bereich des nicht behandelten Stoffes entfallen. Die entsprechenden Werte der Variablen wurden mit l bezeichnet. Wir betonten, daß diese Werte nicht direkt beobachtbar sind. Unter der Voraussetzung, daß das Subjekt den Anteil π des Lehrstoffes nicht beherrscht hat und daß der Test n Aufforderungen enthält, ist die zufällige Veränderliche Y „a priori" der Binomialwahrscheinlichkeit $P(l \mid \pi)$ (2) gemäß verteilt. Die bedingte Wahrscheinlichkeitsverteilung der direkt beobachtbaren Veränderlichen X unter der Voraussetzung, daß l Fragen auf den Bereich des nicht beherrschten Stoffes entfallen sind, wurde durch $P(k \mid l)$ in (3) angegeben. Hierbei spielt l die Rolle des Verteilungsparameters. Die Randverteilung der Veränderlichen X wurde durch die Binomialwahrscheinlichkeit (4) eingeführt. Mittels der Bayesschen Formel [3,4]

$$(9) \qquad P(l\,|\,k) = \frac{P(l)\,P(k\,|\,l)}{P(k)}$$

kann die Wahrscheinlichkeitsverteilung der Zufallsvariablen „a posteriori" abgeleitet werden. Setzen wir die entsprechenden Ausdrücke ein, so bekommt man die Binomialverteilung

$$(10) \qquad P(l|k) = \binom{n-k}{l-k}\left(\frac{\pi}{\pi + q(1-\pi)}\right)^{l-k}\left(\frac{q(1-\pi)}{\pi + q(1-\pi)}\right)^{n-l}$$

$(l = k,\ k+1, \ldots, n)$
$(k = 0, 1, \ldots, n)$

Aus (10) kann der Mittelwert der Variablen Y in Abhängigkeit von X, d.h. die Regressionsfunktion von Y bezüglich X, abgeleitet werden. Sie stellt eine Schätzung von Y bei einem beobachteten Wert von X dar

$$(11) \qquad E(l|k) = k\,\frac{q(1-\pi)}{\pi + q(1-\pi)} + n\,\frac{\pi}{\pi + q(1-\pi)}$$

Der Ausdruck ist linear in k.

Für k = 0 (keine falsche Antwort im Test) kann man erwarten, daß für das Subjekt in Wirklichkeit

$$E(l|k=0) = n\,\frac{\pi}{\pi + q(1-\pi)}$$

Testfragen aus dem Bereich des nicht beherrschten Stoffes stammen.

Für k = n (alle Testfragen wurden falsch beantwortet) kann man $E(l|k = n) = n$ erwarten.

Die Bayessche Schätzung hängt vom Zufallsergebnis k ab. Deshalb ist es zweckmäßig, ihren Mittelwert und ihre Streuung zu bestimmen. Es gilt dabei

$$E_k\{E_l(l|k)\} = n\pi,$$

d.h. der Mittelwert der Bayesschen Schätzung (11) ist gleich dem Mittelwert der unbedingten Verteilung (2).

Der Ausdruck

$$(12) \quad D_k\{E_l(l|k)\} = n\pi(1-\pi)\left\{1 - \frac{1}{\pi + q(1-\pi)}\right\}$$

stellt die Variabilität der Regressionsfunktion $E\ (l \mid k)$ um den gemeinsamen Mittelwert $E(l)$ dar. Die totale Streuung der Zufallsveränderlichen Y kann in zwei Komponenten geteilt werden. Die erste ist durch (12) gegeben, die andere ist der Mittelwert der Reststreuung $D\ (l \mid k)$ der bedingten Wahrscheinlichkeitsverteilung (10). Weil diese eine Binomialverteilung ist, wird die Streuung durch

$$(13) \quad D(l|k) = (n-k)\frac{\pi}{\pi + q(1-\pi)}\frac{q(1-\pi)}{\pi + q(1-\pi)}$$

angegeben. Für ihren Mittelwert erhält man

$$(14) \quad E\{D_l(l|k)\} = n\pi(1-\pi)\frac{1}{\pi + q(1-\pi)}.$$

Das Verhältnis der beiden Streuungskomponenten hängt von π und q ab, nicht aber vom Testumfang n. Die Reststreuung dient dabei als Maß für die Wirkung des Zufalls, daß bei der Auswahl einer Antwortvorgabe eine richtige Antwort nicht nur auf Grund vorhandenen Wissens, sondern auch durch Raten erreicht werden kann. Dieser Anteil sinkt mit zunehmender Anzahl der Antwortvorgaben von einem Minimum $1/(2-\pi)$ für $q = 2$ bis zu 0 für $q = \infty$.

11. *Empirische Bayessche Schätzung des Parameters π*

Setzen wir voraus, daß der gegebene didaktische Test in einer Gruppe von N Subjekten angewandt wurde und daß eine „a priori" - Wahrscheinlichkeitsverteilung $F(\pi)$ in einer Population der Subjekte existiere. Die Randverteilung der Anzahl der Fehler k in einem didaktischen Test mit n Fragen und q Antwortvorgaben für jede Frage wird dann durch das Integral

$$(15) \quad P(k) = \int_0^1 P(k|\pi)\, dF(\pi)$$

gegeben, wobei $P(k|\pi)$ die bedingte Verteilung (4) ist. Wichtig ist, daß sich (15) durch die empirische Häufigkeitsverteilung $N(k)/N$, $k = 0,1,\ldots,n$, in der Gruppe von N Subjekten abschätzen läßt.

Die „a posteriori" - Verteilung des Anteils π des nicht beherrschten Lehrstoffes in der Untermenge der Subjekte, die im didaktischen Test k Fehler gemacht haben, kann dann in die Form

$$(16) \quad dF(\pi|k) = \frac{P(k|\pi)\ dF(\pi)}{P(k)} = \frac{1}{P(k)}\binom{n}{k}\left(\pi\frac{q-1}{q}\right)^k\left(1-\pi\frac{q-1}{q}\right)^{n-k} dF(\pi)$$

gebracht werden.

Diese Formel ermöglicht den Mittelwert der Parameterfunktion $\pi(q-1) \mid \{q-\pi(q-1)\}$ zu berechnen

$$(17) \quad E\left\{\frac{\pi(q-1)}{q-\pi(q-1)} \mid k\right\} = \frac{k+1}{n-k}\frac{P(k+1)}{P(k)}.$$

Nimmt man jetzt die relativen Häufigkeiten $N(k)/N$ anstatt $P(k)$, so bekommt man die Beziehung

$$(18) \quad E\left\{\frac{\pi(q-1)}{q-\pi(q-1)} \mid k\right\} = \frac{k+1}{n-k}\frac{N(k+1)}{N(k)},$$

die für jeden beobachteten k-Wert ($k = 0,1,\ldots, n$-1) die empirische Bayessche Schätzung $\hat{\pi}(k)$ von π bietet [1,2]

$$(19) \quad \hat{\pi}(k) = \frac{q}{q-1}\frac{\frac{k+1}{n-k}\frac{N(k+1)}{N(k)}}{1+\frac{k+1}{n-k}\frac{N(k+1)}{N(k)}}$$

Durch die Bayessche Schätzung wird ein Prinzip quantifiziert, daß ein Lehrer bei der Bewertung eines Schülers mehr intuitiv anwendet: Er bezieht außer der aktuell gebotenen Leistung auch dessen frühere Ergebnisse und die Leistungen der anderen Kollektivmitglieder in die Bewertung ein.

Schrifttum

Komenda, S., J. Klementa: Ein wahrscheinlichkeitstheoretisches Modell für einen didaktischen Text und seine statistische Analyse. Wiss. Zeit., Halle XVI (1978), S. 3-9

Komenda, S., J. Klementa: Analýza náhodného v pedagogickém experimentu a praxi. SPN Praha, 1981

Komenda, S.: Bayes rule in the relation between students knowledge and response. Referate des 20. Int. Symp. „Ingenieurpädagogik 91", S. 114-119, Dresden 1991
Sachs, L.: Applied Statistics. A Handbook of Techniques. 2nd Ed. Springer-Verlag 1984
Thorndike, R.L. (Ed): Educational Measurement. 2nd Ed. New York 1971

Eingegangen 1994-12-20

Anschrift des Verfassers: Prof. RNDr. Stanislav Komenda, Dr. Sc., Hněvotínská 3, CZ-77515 Olomouc

School - achievement test from the statistical viewpoint (Summary)
When the knowledge the subject has about the topic under consideration is to be assessed, three levels are reasonably recognized: (a) the ratio π ($0\leq\pi\leq1$) of the topic under examination unmastered by the subject, (b) the number l among the n items of the test the subject does not know, and (c) the number k($0\leq k\leq l$) of the test items answered incorrectly by the subject (which is expected to be lower in comparison with l due to the non-zero chance for the ignoring subject to reach the correct response by guessing).
The methods of inductive statistics make it possible to estimate numerical values of the test parameters l and π as well as by means of the immediately observable number of false responses k. The method of maximum likelihood, inverse probability and the empirical Bayes procedure were applied as well. The basic properties of the estimates thus obtained were derived and the respective formulas found. The key points of the procedures under consideration were demonstrated in graphical form.

Lernejsukcesa testo el statistika vidpunkto (resumo)
Se la scio, kiun havas la subjeko pri la konsiderata objekto, devas esti stimata, oni racie akceptas tri nivelojn: (a) la procentaĵon π ($0\leq\pi\leq1$) de la ekzamenata temo nekonata de la subjekto, (b) la nombron l el la n testeroj, kiujn la subjekto ne konas, kaj (c) la nombron k ($0\leq k\leq l$) de la testeroj malĝuste responditaj de la subjekto (kiun oni supozas malpli granda kompare kun l pro la ne-nulo ŝanco por la nescianta subjekto atingi la ĝustan respondon per divenado).
La metodoj de induktiva statistiko ebligas stimi numerajn valorojn de la testparametroj l kaj π samkiel helpe de la direkte observebla nombro de malĝustaj respondoj k. La metodo de maksimuma probableco, inversa probablo kaj la empiria procedo de Bayes estas ankaŭ aplikataj. La bazaj ecoj de la stimataĵoj tiel ricevitaj estis derivitaj kaj korespondantaj formuloj estis trovitaj. La ŝlosilpunktoj de la konsiderataj procedoj estis montrataj en formo de grafikaĵo.

grkg / Humankybernetik
Band 38 · Heft 3 (1997)
Akademia Libroservo / IfK

A Mathematical Analysis of Problem Situations - Structure and MAPLE

by Josef BRODY, Montreal (CND) and Milan HEJNÝ, Prague (CZ)

from Department of Mathematics and Statistics, Concordia University, Montreal and from Department of Mathematics and Didactics of Mathematics, Faculty of Education, Charles University, Prague

Introduction

In order to make this article self-contained, we refer the reader to our report (Brody, Hejný and Rosenfield, 1996), where more elaborated views including details and proofs, could be found in. In Part 1 (Brody, Hejný, 1997) we have introduced the following Problem Situation *PS*:

Two men went for a walk. They left their hotel shortly after their lunch at 2 P.M. and walked to the top of a hill. They returned to their hotel using the same route and taking no longer than they did on the trip uphill. They wanted to be in the hotel by 8 P.M.

In this *PS* we identified the following objects:

distance in km: one-way (d) - up or down, total (dt);

times in hours: up (tu),down (td), total (tt), difference (t) - of tu and td;

speed in km/hour: up (su), down (sd), overall (so);

coefficients (with no measurements): (ct) - ratio of tu and td.

Based on these we defined the *object set* $\mathcal{O} = \{d, dt, tu, td, tt, t, su, sd, so, cd\}$, and a *free situation* $Sf = [d, dt, tu, td, tt, t, su, sd, so, ct]$. To a free situation we assign $\mathbf{x} \in \mathbb{R}_{10}$ denoted by $Sf := \mathbf{x}$ and get a *concrete situation* $Sc = (Sf := \mathbf{x})$. The *set of* (such all) *concrete situations* is denoted by $\mathcal{O}_0$.

Since not every Sc satisfies the *PS* we defined two kinds of relations:

1) Rc - the set of all *comparative relations*. All Sc satisfying Rc[1] are meaningful. The set Ω is the set of all meaningful assignments $\mathbf{x} \in \mathbb{R}_{10}$;

[1] A concrete situation Sc satisfies a set of relations R means that Sc satisfies every relation $r \in R$.

2) Rd - the set of all *determinative relations*. All Sc satisfying *Rd* are *feasible*. The set $\sum_{Rd}$ is the set of all feasible assignments $\mathbf{x} \in \Omega$.

Any relation r is a mapping: $\mathbb{R}_{10} \to \{\mathbb{T}, \mathbb{F}\}$ (where $\mathbb{T}$ stands for true and $\mathbb{F}$ for false). The *characteristic set*: $\Sigma_r = \{\mathbf{x} \in \mathbb{R}_{10} : r(\mathbf{x}) = \mathbb{T}\}$ is called the *domain* of the relation r. To every $\mathrm{x} \in \Omega$ and $i \in \{1,...,10\}$ we define the *i*-projection of the domain $\sum_r: \sum_r(\mathrm{x};i) = \{y \in \Omega : [x_1,\ldots x_{i-1},y,x_{i+1},\ldots x_{10}] \in \sum_r\}$. We call relation r a *determinative relation* on Ω iff[2] for every $\mathbf{x} \in \Omega$ the *i*-projection of $\sum_r$ is either a singleton, or the *i*-projection of $\Omega : \Omega(\mathrm{x};i) = \{y \in \mathbb{R}: [x_1,\ldots x_{i-1},y,x_{i+1},\ldots,x_{10}] \in \Omega\}$. If the *i*-projection of $\sum_r$ is a singleton we call the index i and the object o_i (the *i*- th component of *Sf*) *dependent on r*. If the *i*-projection of $\sum_r$ is $\Omega(\mathbf{x};i)$ we call i and o_i *independent of r*. The number of objects dependent on r is called the *size* of r , denoted as: $\sigma_r = |r|$ and the set of all dependent objects on r is denoted by $\mathcal{O}_r$.

Let $r_1\ldots,r_m$ be determinative relations on Ω. Then the domain of $S = \{r_1,\ldots,r_m\}$ is given as $\sum_S = \bigcap_{i=1}^{m} \sum_{r_i}$. A determinative relation r is said to be *a logical consequence* of S if $\sum_S \subset \sum_r$. The set of all logical consequences of S is called *the envelope of S* and denoted by S_{env} = env(S). We call S a *dependent set* if there is a relation $r_i \in S : \bigcap_{j \neq i} \sum_{r_j} \subset \sum_{r_i}$ (i.e. r_i is a logical consequence of $S - \{r_i\}$. In the opposite case S is said to be an *independent set*.

Let S be a determinative set of relations. An independent set $S_0 \subset S : \sum_S = \sum_{S_0}$ is called a *basis* of S. If S_1 is another basis of S then it was shown that $|S_0| = |S_1|$. This common number of relations is called *rank of S* and denoted by rank (S) which also defines the dimension of S_{env} as dim $(S_{env}) = \delta$.

All the above notions and ideas concerning our "two-knights' walk" problem situation can be generalized to a general Problem Situation *PS*, with $\mathcal{O} = \{o_1,...,o_n\}$, $Sf = [o_1,\ldots,o_n]$, $\mathrm{x} = [\mathrm{x}_1,\ldots,\mathrm{x}_n] \in \mathbb{R}_n$, $Sc = (Sf := \mathrm{x})$, $\Omega \subset \mathbb{R}_n$, $\sum_r(\mathrm{x};i) \subset \Omega(\mathrm{x};i) \subset \mathbb{R}$, etc.

1. *Structure of a problem situation*

In this article we consider only such determinative sets of relations S which have a non-empty domain. Therefore we have rank (S) $\leq |\mathcal{O}|$ i.e. the number of relations in any of its bases cannot exceed the number of objects in $\mathcal{O}$. In this section we shall discuss an *algebraic structure* of infinite set env(S) = *Rd* of all determinative relations which are logical consequences of S.

[2] *iff* means *if* and only *if*.

Definition 1.1: Let S be a determinative set of relations. Then we define *dimension of Rd* denoted by dim(*Rd*) as the rank(S) ⋈◁⋈

In our example the set of objects $\mathcal{O}$ = {d ,*dt, tu, td, tt, t,su, sd, so, ct*} has ten elements, therefore any determinative set of relations S has rank(S) $\leq$ 10. Let us introduce all intrinsic properties of our problem situation *PS* and define the following set of determinative relations $S = \{s_1,\ldots,s_7\}$[3] where $s_1 = (dt = 2d)$, $s_2 = (tt = tu + td)$, $s_3 = (t = tu - td)$, $s_4 = (su = d/tu)$, $s_5 = (sd = d/td)$, $s_6 = (so = 2d/tt)$, $s_7 = (ct = tu/td)$. We can see that each of these relations s_i "introduces" one object o_i (on the left side of its defining equation) which depends only on s_i, i.e. does not depend on any other relations s_j, where $i \neq j$. This property makes these (seven) relations obviously independent. To underline this "simplicity" we state:

Definition 1.2: Let $S = \{s_1,...,s_p\}$ be a determinative set of relations such that *Rd* = env(S). Let $\mathcal{O}_S$ be the set of all objects dependent on S an for each $i = 1,...,p$ exists an object $o_i \in \mathcal{O}_S$ such that o_i is dependent on s_i but does not depend on any (other) s_j, where $i \neq j$. Then S is called a *canonical basis* of *Rd* ⋈◁⋈.

Definition 1.3: Let $\mathcal{O}(PS)$ be a set of all objects in a problem situation *PS*. Let T_{PS} with domain $\sum_{PS}$, be a determinative set of relations which captures all **intrinsic** properties of the problem situation *PS*. Let the set $\mathcal{O}_{PS} \subset \mathcal{O}(PS)$ be the set of all objects dependent on T_{PS}.[4] Then we call the couple $S_{PS} = [\mathcal{O}_{PS}, \Sigma_{PS}]$ the *structure of the problem situation PS*. The determinative set of relations R_{PS} = env(T_{PS}) is called the *relation space of the problem situation PS.* We call the dimension of the relation space the *dimension of the problem situation PS* and denote it by dim(*PS*). The number of objects in $\mathcal{O}_{PS}$ is then called the *size of problem situation PS* and denoted by s(*PS*) = $\sigma = |\mathcal{O}_{PS}|$ ⋈◁⋈.

Comment: Obviously we have dim (*PS*) = dim (R_{PS}) = rank(T_{PS}) ⋈◁⋈.

In our example we have $\mathcal{O}_{PS} = \mathcal{O}$, $s(PS) = 10$, $T_{PS} = S$, R_{PS} = env (S), and dim (*PS*) = 7. With each of the above relations we can associate a function; for example with the relation s_2 we can associate the function $f_{s_2} : f_{s_2}(\mathbf{x}) = x_3 + x_4 - x_5$, and see that $\mathbf{x} \in \sum_{s_2}$ iff $f_{s_2}(\mathbf{x}) = 0$ Generally we have:

Definition 1.4: A relation s is *associated* with a function f_s: $\mathbb{R}_n \to \mathbb{R}$, denoted as $s \leftrightarrow f_s$, if $\mathrm{x} \in \sum_s$ iff $f_s(\mathbf{x}) = 0$ ⋈◁⋈.

Let us discuss our example with the set of all meaningful concrete situations $\Omega \subset \mathbb{R}_{10}$. The determinative relation (above) s_2 is associated with f_{s_2} $(s_2 \leftrightarrow f_{s_2})$, where $f_{s_2} = x_3 + x_4 - x_5$, defines the set of all objects dependent on s_2:$\mathcal{O}_{s_2}$ where $\mathcal{O}_{s_2}$ = {*tu, td, tt)* $\subset \mathcal{O}$. Each of the objects in $\mathcal{O}_{s_2}$ depends on the other two, i.e. $x_3 + x_4 - x_5 = 0$ iff: $x_3 = -x_4 + x_5 (= f_{s_2;3}(x_4,x_5)), x_4 = -x_3 + x_5 (= f_{s_2;4}(x_2,x_5)), x_5 = x_3 + x_4$ $(= f_{s_2;5}(x_3,x_4))$. We shall write $f_{s_2}(\mathbf{x})$ instead of $f_{s_2}(x_3,x_4,x_5)$, since the other seven

[3] In part one we used rd_i instead of simpler notation s_i we prefer to use here.

[4] This means that for each $o \in \mathcal{O}_{PS}$ exists a relation $r \in T_{PS}$ such that o depends on r, i.e. $o \in \mathcal{O}_r$.

values are assigned to the seven independent objects of the relation s_2. We can therefore write $f_{s_2;3}(\mathbf{x}_{;3})$ instead of $f_{s_2;3}(x_4,x_5), f_{s_2;4}(\mathbf{x}_{;4})$ instead of $f_{s_2;4}(x_2,x_5)$, and $f_{s_2;5}(\mathbf{x}_{;5})$ instead of $f_{s_2;5}(x_3,x_4)$.[5] We shall now state the following definition:

Definition 1.5: Let $i \in \{1,...,n\}, \Omega \subset \mathbb{R}_n$ and $\Omega_i = \bigcup_{\mathbf{x}\in\Omega}\{\mathbf{x}_{;i}\} \subset \mathbb{R}_{n-1}$. A function $f_s : \Omega \to \mathbb{R}$, associated with a determinative relation s, is called *i-solvable on* Ω if there exists $f_{s;i} : \Omega_i \to \mathbf{R}$ that for every $\mathbf{x} = [x_1,...,x_n] \in \Omega$: $f_s(\mathbf{x}) = 0$ iff $x_i = f_{s;i}(\mathbf{x}_{;i})$ ⋈◁⋈.

In the above example the function f_{s_2} is i- solvable on $\mathbf{R}_{10}$ for i = 3,4,5. However, we are only interested in its solvability for meaningful concrete situations, i.e. on the set Ω. This places the point $\mathbf{x}_{;3}$ into Ω_3 (= $\bigcup_{\mathbf{x}\in\Omega}\{\mathbf{x}_{;i}\} \subset \mathbb{R}_9$), $\mathbf{x}_{;4}$ into Ω_4, and $\mathbf{x}_{;5}$ into Ω_5. More generally:

Theorem 1.1: Let Ω be the set if all meaningful concrete situations of a problem situation PS and r a relation associated with F_r. Then for every i : $o_i \in \mathcal{O}_r$ the function f_r is i- solvable on Ω iff r is a determinative relation (in Ω) ⋈◁⋈

Comment: Let $r_1,...,r_k$ be determinative relations such that $r_i \leftrightarrow f_{r_i}$. Let g : $\mathbb{R}_k \to \mathbb{R}$ be such that: $g(\mathbf{0}) = 0$. Then we denote the relation r associated with $g(f_{r_1},...,f_{r_k})$ by $g(r_1,...,r_k)$ ⋈◁⋈

Comment: From theorem 1.1 it follows that the relation $r = g(r_1,...,r_k)$ is determinative on Ω iff the function $g(f_{r_1},...,f_{r_k})$ is i- solvable on Ω for every i: $o_i \in \mathcal{O}_r$ ⋈◁⋈.

In our example we can se that the relations $s_2 \leftrightarrow f_{s_2} : f_{s_2}(\mathbf{x}) = x_3 + x_4 - x_5$ and $s_3 \leftrightarrow f_{s_3} : f_{s_3}(\mathbf{x}) = x_3 - x_4 - x_6$. From these two relations we can now "create" (infinite number of) new determinative relations (s) like $r = g(f_{s_2}, f_{s_3})$[6] $= s_2 + s_3 \leftrightarrow 2x_3 - x_5 - x_6$. We can also see that r is a logical consequence of the determinative set of relations $S = \{s_2, s_3\}$. More generally we have:

Theorem 1.2: Let $r_1,...,r_k$ be determinative relations on Ω such that $r_i \leftrightarrow f_{r_i}$. Let g: $\mathbb{R}_k \to \mathbb{R}$ be such that $g(\mathbf{0}) = 0$. Let for all dependent objects o_i on relation $r = g(r_1,...,r_k)$ the function $g(f_{r_1},...,f_{r_k})$ be i- solvable on Ω. Then r is a determinative relation which is a logical consequence of the determinative set of relations $S = \{r_1,...r_k\}$ ⋈◁⋈.

2. *Solving processes*

Since in our example we have dim(PS) = rank(S) = 7, there must exist a triple of "independent" objects $\mathcal{O}'' = \{o_{i_1}, o_{i_2}, o_{i_3}\} \subset \mathcal{O}_{PS}$ such that any assignment of the other seven objects $\mathcal{O}' = \{o_{j_1},...,o_{j_7}\} \subset \mathcal{O}_{PS}$ depend on the assignment of the three in $\mathcal{O}''$. Therefore it

[5] Realize tha $\mathbf{x}_{;i} = [x_1,...,x_{i-1},x_{i+1},...,x_{10}] \in \mathbf{R}_9$ is the *i-th* projection of $\mathbf{x}$. Here i=3,4,5

[6] The function $g : \mathbf{R}_2 \to \mathbf{R}$ is $g(x,y) = x + y$, i.e. $g(0,0) = 0$ and is 1 (and 2)-solvable since $g(x,y) = 0$ iff $x = f_1(y) = -y$ and $y = f_2(x) = -x$.

follows that the $(j_1,...,j_7)$; 7- projection of Σ_S: $\Sigma_S(\mathbf{x}; j_1,...,j_7)$ is a singleton for every feasible $\mathbf{x} \in \Sigma_S$. We refer to such three objects as "independent" (of each other). This leads us to the following notions:

Definition 2.1: Let PS be a problem situation with the structure $S_{PS} = [\mathcal{O}_{PS}, \Sigma_{PS}]$, relations space Rd, size $s(PS) = \sigma$ and dimension dim(PS) = δ. We call a set $\mathcal{O}'' \subset \mathcal{O}_{PS}$ a *dependent set of objects within the structure* S_{PS} iff there exists a relation $r \in Rd$ $\mathcal{O}_r \subset \mathcal{O}''$. In the opposite case we call the set $\mathcal{O}''$ an *independent set of objects within the structure* S_{PS} ⋈◁⋈.

Then it is obvious that:

Theorem 2.1: Let $\mathcal{O}_1 \subset \mathcal{O}'' \subset \mathcal{O}_2 \subset \mathcal{O}_{PS}$. If $\mathcal{O}''$ is a dependent set of objects in structure S_{PS} then so is $\mathcal{O}_2$. If $\mathcal{O}''$ is an independent set of objects in structure S_{PS} then so is $\mathcal{O}_1$ ⋈◁⋈.

Definition 2.2: Let PS be a problem situation with the structure $S_{PS} = [\mathcal{O}_{PS}, \Sigma_{PS}]$, size $s(PS) = \sigma$, and dimension dim(PS) = δ. Let $\varphi = \sigma - \delta$ and let $\mathcal{O}'' = \{o_{j_1},...,o_{j_\varphi}\}$ be an independent set of objects within the structure S_{PS}. Then the $(i_1,...,i_\delta)$; δ-projection of Σ_{PS}: $\Sigma_{PS}(\mathrm{x}; i_1,...,i_\delta)$ is a singleton for every $\mathrm{x} \in \mathbb{R}$. This projection is associated with the set $\mathcal{O}' = \{o_{i_1},...,o_{i_\delta}\}$ called *a basic object set* and object in $\mathcal{O}'$ is called *a basic object*. The complementary set $\mathcal{O}'' = \mathcal{O}_{PS} - \mathcal{O}'$ is called a *parameter set* and each its element a *non-basic object*. The feasible values $x_{j_1},...,x_{j_\varphi}$ assigned to non-basic objects in order to get a feasible partial concrete situation $Spc_{i_1,...,i_\delta} = (Sf := \mathrm{x}_{;\{o_{i_1},...,o_{i_\delta}\}})$[7] are called *parameters*. The size φ of the parameter set is called the *degree of freedom* df(PS) = φ *of the problem situation PS* ⋈◁⋈.

Comment: The degree of freedom φ gives the number of a"maximum sized"[8] independent set of objects $\mathcal{O}''$ within the structure S_{PS}, while the dimension δ establishes the number of *independent relations* in T_{PS}. Since $\varphi + \delta = \sigma$ we consider the two sets ($\mathcal{O}_{PS}$ and T_{PS}) *mutually dual*. The duality is actually between the (infinite) *set of feasible concrete situations* $\mathcal{O}_0$ *over* $\mathcal{O}_{PS}$ and the (infinite) *set of all determinative relations Rd* over T_{PS} ⋈◁⋈.

Definition 2.3: Let $\mathcal{O}'$ $\{o_{i_1,...,o_{i_\delta}}\}$ be such that the $(i_1,...,i_\delta)$;δ-projection $\Sigma_{PS}(\mathbf{x}; i_1,...,i_\delta)$ is a singleton for each feasible $\mathbf{x}$. Then the related projection vector function $\mathbf{f}_{PS;i_1,...,i_\delta} = \left[f_{PS;i_1},...,f_{PS;i_\delta}\right]$ is called a *solution with respect to* $\mathcal{O}'$, and the relations $r_{i_1},...,r_{i_\delta}$ given by equations:

$$x_{i_1} = f_{PS;i_1}(\mathrm{x}_{;i_1,..i_\delta}),...,x_{i_\delta} = f_{PS;i_\delta}(x_{;i_1,..i_\delta})$$

[7] Note that $\mathrm{X}_{;\{i_1,...,i_\delta\}} = [x_1,...,x_{i_1-1},-,x_{i_1+1},...,x_{i_\delta-1},-,x_{i_\delta+1},...,x_n]$ where $i_1,...,i_\delta$ indicate the δ blank positions of $\mathrm{X}_{;\{i_1,...,i_\delta\}}$.

[8] That means that each set $\mathcal{O}_2 \supset \mathcal{O}''$ such that $\mathcal{O}_2 \neq \mathcal{O}''$ is already a dependent set of objects in structure S_{PS}.

are called a *canonical basis with respect to* $\mathcal{O}''$[9] The process where a parameter set $\mathcal{O}''$ is the input and its canonical basis is the output is called a *solving process* ⊳⊲⊳⊲.

Comment: There are many canonical bases in every problem situation *PS*. However, the maximum number of them is limited by the combinatorial number $\binom{\sigma}{\varphi}$. Some φ-tuples may be dependent and therefore the binomial number $\binom{\sigma}{\varphi}$ is generally only the upper bound on the number of canonical bases ⊳⊲⊳⊲.

In our example we have degree of freedom $\varphi = df(PS) = 3$, dimension $\delta = \dim(PS) = 7$ and size $\sigma = s(PS) = 10$, so the upper bound on the number of canonical bases is $\binom{10}{3} = 120$. However, some triples are dependent e.g. $\mathcal{O}'' = \{tu, td, tt\}$, since the relation s_1: $tu + td = tt$ is defined on $\mathcal{O}''$. Actually there are 28 dependent triples. All the other 92 triples are independent. This means that there are 92 canonical bases in our problem situation *PS*. The corresponding solution strategies can be found by means of mathematical software, e.g. MAPLE. One of the 92 independent triples is $\mathcal{O}'' = \{su, sd, d\}$. It gives the following canonical basis:

$$\begin{array}{ll} r_1:\ tu = d/su & r_2:\ tt = d(sd+su)/sd.su \\ r_3:\ td = d/sd & r_4:\ t = d(sd-su)/sd.su \\ r_5:\ dt = 2d & r_6:\ sa = 2su.sd/(sd+su) \\ r_7:\ sc = sd/su & \end{array}$$

All such independent parameter sets $\mathcal{O}''$ represent 92 open problems since each such independent triple generates a canonical basis which in turn defines an assignment for each of the remaining seven non-basic object. Each such assignment represents an extension of a partial concrete situation

$$Spc_{\mathcal{O}'}:\ Spc_{\mathcal{O}'} = (Sf_{PS} := \mathrm{x}_{;i_1,\ldots,i_7}) = (Sf_{PS} := [x_1,\ldots,x_{i_1-1},-,x_{i_1+1},\ldots,x_{i_7-1},-,x_{i_7-1},\ldots,x_{10}])$$

where $\mathcal{O}'$ is the basic set, to a (full) concrete situation: $Sc = (Sf := \mathbf{x})$. This is why the MAPLE-solving process for each independent triple ends with the allocation of the corresponding canonical basis.

3. *Experimenting with MAPLE software*

The choice of MAPLE was bases on the capability of the MAPLE function *solve*. The procedure (program) **REL(N;M;P)** contains three parameters which have to be chosen by the user. These parameters are natural numbers $(N \le \sigma, M \le P \le \binom{\sigma}{N})$. After execution of this procedure, MAPLE generates a list of all *N*- tuples (not yet printed). The order of these is made up by the package from the MAPLE library COMBINAT, CHOOSE. The user's choice of the parameters *M, P* makes MAPLE to list all of these *N*-tuples starting with the *M–th* and terminating with the *P–th*. Let us denote $\{o_{i_1},\ldots,o_{i_N}\} =$

[9] Remember that the set of parameters $\mathcal{O}''$ is the complement to $\mathcal{O}'$ in $\mathcal{O}_{PS}$ i.e. $\mathcal{O}'' = \mathcal{O}_{PS} - \mathcal{O}'$.

$\mathcal{O}'' \subset \mathcal{O}_{PS}$. If $\mathcal{O}''$ defines a dependent set of objects, no solution is presented[10]. If $\mathcal{O}''$ defines an independent set of objects, the canonical basis with respect to $\mathcal{O}''$ is presented.

Let us assume that the program was informed about the object set $\mathcal{O}_{PS}$ and the relation set T_{PS}. Let us further assume that the problem Situation *PS* has size σ, dimension δ, and the degree of freedom φ. Each student starts with filling inside of the procedure the particular set of objects $\mathcal{O}_{PS}$, the determinative set of relations T_{PS} and the size σ. When a user is equipped with experiences which enable him/her to penetrate into difficult concepts such as independence and duality, then the above procedure, offers the following possible situations:

1. If $N < \varphi$, then the procedure (randomly) adds φ - N objects to each *N*- tuple, which indicates that each partial concrete situation is *under-determined.*
2. If $N = \varphi$, each independent set of objects $\mathcal{O}''$ produces the canonical basis with respect to $\mathcal{O}''$ and a dependent set $\mathcal{O}''$ does not produce anything. This enables the user to experience notions of *dependent objects, independent objects degree of freedom of the* problem situation φ, and its dimension δ.
3. If $N > \varphi$, each object set $\mathcal{O}''$ produces no solution, which indicates, as mentioned above, that all N - tuples are dependent and each corresponding partial situation is *over-determined.*
4. Comparison of $|T_{PS}|$ and φ with the size σ lets the user know whether the relation set T_{PS} is independent or dependent. In case that observed $|T_{PS}| + \varphi = \sigma$ the relation set is an independent set, i.e. a basic one, and $|T_{PS}| = \dim(PS) = \delta$. If $|T_{PS}| + \varphi > \sigma$, then $|T_{PS}| > \dim(PS) = \delta$ and the relation set is a dependent set and contains a basic subset $S \subset T_{PS}$ such that $|S| = \delta = \sigma - \varphi$ (already observed in case two above).

In cases one to three the student works with the set of objects $\mathcal{O}_{PS}$, while in case four he/she can observe the relation space *Rd* and its duality with the space of all concrete situations $\mathcal{O}_0$. Besides this experimental view into the structure of the problem situation *PS* one can also realize the expansions of a large quantity of determined partial concrete situations $Spc_{\mathcal{O}'}$ into a full concrete situation *Sc*, since each canonical basis consists of all necessary formulas and the only remaining step is their use for calculation.

Conclusion

1) We do believe that profound difficulties in Linear or Abstract Algebra for university students are in the traditional lecturing theoretical approach to these subjects and lack of concrete models.

2) According to our teaching experiences the described analysis of a problem situation enables students

[10] For example a dependent triple $\{tu, td, tt\}$.

- to penetrate into notions like *relations, (in)dependence, basis, duality,* etc. and applying this knowledge to MAPLE
- to create several different and well elaborated concrete situations which equip students with experiences necessary
- to get insight into theoretical concepts.

3) If in addition the student is a future teacher of Mathematics this knowhow enables him/her to individualize his/her teaching approach.

Acknowledgment: This paper was supported through the grant GACR 403/96/1186.

References

Brody, J., Hejný, M.: *A Mathematical Analysis of Problem Situation, Determinative Relations,* grkg 2/97, 69-76

Brody, J., Hejný, M. , Rosenfield, S.: *A Mathematica Structure for Problem Situations,* Technical Report No. 2/96, August 1996, Department of Mathematics and Statistics, University, Montreal, Canada

Brody, J., Hejný, M. , Rosenfield, S.: *Syntax Phenomenon in the Solving Processes,* Technical Report No. 8/92, October 1992, Department of Mathematics and Statistic, Concordia University, Montreal, Canada

Received 1996-12-29

Address of the authors: Prof.Dr. Josef Brody, 5795 Sir Walter Scott, App. 804, Montreal/Quebec, H4W2TT, Canada, Prof.Dr. M. Hejný, Nušlova 2272, CZ-15800 Praha 5

Matematika analizo de problemaj situacioj. Strukturo kaj MAPLE.(Resumo)

Ni evoluigis aliron al la problemsolvado, uzeblan en analizado de problemsituacio helpe de konceptoj kiel objektoj, rilatoj, projekcioj, sendependeco, bazo, dimensio ktp. En tiu ĉi artikolo ni volas pridiskuti strukturon kaj dualecon de la spaco de ĉiuj konkretaj situacioj $\mathcal{O}_0$ kaj la spacon de ĉiuj determinaj rilatoj *Rd*. MAPLE estas uzata por listigi ĉiujn eblajn kanonikajn bazojn de la diskutata problemsituacio. Ni montras, kiel tia listigo helpas la analizon de ambaŭ spacoj $\mathcal{O}_0$ kaj *Rd*. Studentoj havantaj spertojn ebligantajn penetri en malfacilajn konceptojn kiel sendependeco kaj dualeco, kiuj estas fundamentaj en subjektoj kiel lineara aŭ abstrakta algebro, poste uzas MAPLE por eksperimenti en la donita problemsituacio. Plue nia studo celas identigi obstaklojn, kiuj kaŭzis, ke tiuj subjektoj estas instruataj je fomala nivelo sen iu ajn serioza kompreno.

grkg / Humankybernetik
Band 38 · Heft 3 (1997)
Akademia Libroservo / IfK

Analyse syntaxique combinant deux formalismes au sein d'un Chart hiérarchique

de Pierre-Olivier EL GUEDJ et Pierre NUGUES, Caen (F)

Membres du laboratoire GREYC, commun à l'Institut des Sciences de la Matière et du Rayonnement et à l'Université de Caen, Caen (F).

1 *Introduction*

Cet article décrit un analyseur syntaxique utilisant deux méthodes d'analyses complémentaires (El Guedj 1996; El Guedj & Nugues 1996; El Guedj & Nugues 1994). Lorsqu'on considère la langue française, on peut distinguer deux niveaux de structures syntaxiques. Le premier niveau correspond à des structures dont l'ordre des mots est très précisément défini: il s'agit de ce qu'on appelle communément les groupes syntaxiques (nominal, verbal, prépositionnel). Ces entités syntaxiques sont caractérisées par une certaine rigidité structurelle et l'utilisation d'une grammaire formelle (Chomsky, 1957) s'avère être un moyen adéquat permettant de les décrire puis les repérer au sein de la phrase.

Le second niveau de structure concerne les relations qui existent entre ces groupes syntaxiques. Du fait de la mobilité de certains de ces syntagmes au sein de la phrase, une grammaire de dépendances (Tesnière, 1959; Mel'cuk, 1988; Covington, 1990) se révèle être plus appropriée pour traiter ce second niveau structurel. Une telle grammaire permet d'établir les différentes relations hiérarchiques et fonctionnelles qui existent entre les divers groupes syntagmatiques.

L'utilisation concertée de ces deux paradigmes de représentation syntaxique permet d'optimiser la reconnaissance syntaxique des phrases et d'améliorer la robustesse des analyseurs. De plus, l'application d'une technique similaire à celle des étiqueteurs lexicaux (Chanod & Tapanainen, 1995) à un niveau non plus lexical mais syntaxique permet encore d'améliorer le processus d'analyse des phrases en éliminant les segments ne pouvant s'intégrer avec le reste de la phrase.

Nous avons évalué cet analyseur sur un corpus d'ordres oraux de navigation ainsi que sur une cinquantaine de phrases d'un corpus de comptes rendus médicaux provenant du Centre de traitement anti-cancéreux François Baclesse de Caen.

2 *Architecture du système et principe de fonctionnement*

L'architecture du système d'analyse syntaxique se développe autour d'une structure de *Chart* (Kaplan, 1973; Kay, 1973; Gazdar & Mellish, 1989). Cette structure permet de stocker les hypothèses ainsi que les conclusions partielles de l'analyse en cours. Trois modules agissent directement sur les données contenues dans le Chart: *le module de segmentation syntagmatique, le module d'élagage et le module d'analyse des dépendances* (Figure 1)

Certains de ces modules construisent des hypothèses et les développent pour aboutir à des conclusions partielles. D'autres, au contraire, se chargent d'effectuer un tri parmi ces conclusions afin de favoriser celles qui sont admissibles et de mettre de côté celles présentant un défaut de cohérence ou de vraisemblance.

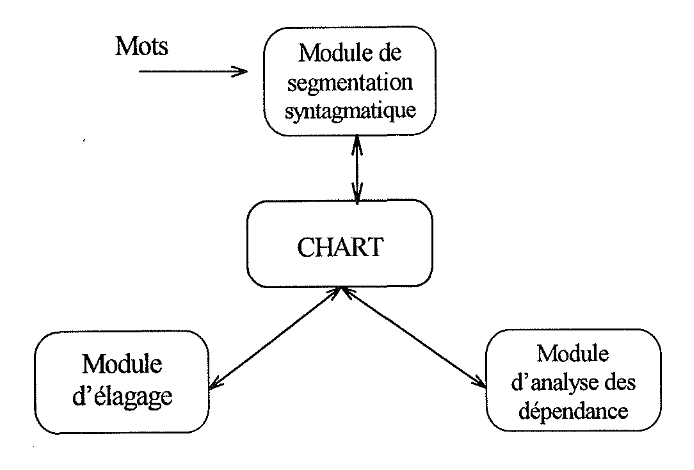


Figure 1: Schéma conceptuel du système

3 *L'organisation hiérarchique du Chart*

Du fait de l'interaction de plusieurs modules sur le Chart, nous avons divisé le Chart en plusieurs niveaux conceptuels à partir desquels les différents modules peuvent intervenir (Figure 2). Notre Chart organise les informations en fonction de la richesse de leur intérêt et du niveau de complexité qu'elles représentent.

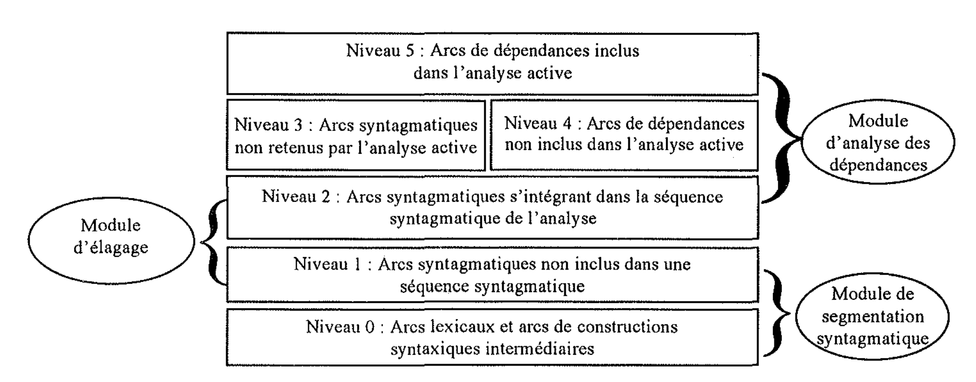


Figure 2: Interaction des trois modules avec les différents niveaux hiérarchiques du Chart

Le niveau 0 est réservé aux arcs portant des hypothèses et des conclusions non syntagmatiques (Arcs lexicaux ou constructions syntaxiques intermédiaires). Le niveau 1 correspond au premier niveau syntagmatique. C'est à ce niveau que le module de segmentation syntagmatique dépose ses conclusions. Le module d'élagage examine l'ensemble de ces conclusions et déplace au niveau 2 les arcs syntagmatiques ayant satisfait aux conditions d'acceptabilité exigée par ce module. Les syntagmes ne remplissant pas les conditions requises sont conservés au niveau 1 qui constitue donc une zone d'archivage des conclusions non retenues pour la suite de la résolution de l'analyse.

Le niveau 5 est réservé aux structures de données intégrant les relations de dépendance et de rattachement mises en place par le module d'analyse des dépendances. Ces relations de dépendance concernent exclusivement les arcs syntagmatiques présents au niveau 2. Le module d'analyse des dépendances peut générer plusieurs analyses possibles de la phrase si elle présente des ambiguïtés au niveau de sa structure. Toutes ces analyses sont consultables mais on ne peut activer qu'une seule analyse à la fois. Elle peut correspondre à un choix établi par le système, l'utilisateur ou un module externe pour une étude ou un traitement particulier (affichage, analyse sémantique...). Les niveaux 3 et 4 servent respectivement à archiver les arcs syntagmatiques et les arcs de dépendance incompatibles ou non utilisés par l'analyse active.

4 *Le module de segmentation syntagmatique*

Le *module de segmentation syntagmatique* repère les groupes syntaxiques présents dans la phrase. Il contient un analyseur syntagmatique dérivé des algorithmes utilisés habituellement dans les systèmes à base de Chart (Earley, 1970; Wirén, 1992). Il accomplit une analyse de gauches à droite. Lors de l'ajout d'un mot au Chart, l'algorithme classique d'Earley considère toute la phrase déjà analysée pour tenter de combiner l'arc du nouveau mot avec ceux déjà présents dans le Chart.

Notre algorithme de segmentation cherche à diminuer cet espace de recherche en ne concentrant l'analyse formelle que sur un segment de la phrase: le syntagme en cours d'analyse (Figure 3). Cette portion de phrase, appelée également *fenêtre d'analyse,* est

délimitée par les deux nœuds (ou *vertex*) du Chart entourant les mots concernés. On nomme le nœud de gauche, le *point d'ancrage*. C'est à partir de lui que le module fonde son analyse formelle. L'analyse syntaxique s'effectuant de gauche à droite au fur et à mesure de la réception des mots, le nœud de droite correspond toujours au dernier nœud du Chart. Un nœud ne peut être un point d'ancrage qu'à deux conditions que nous appellerons les conditions d'ancrage:

- Soit le nœud considéré correspond au premier nœud du Chart;
- Soit il s'agit d'un nœud relié transitivement au nœud 0 par une composition d'arcs inactifs porteurs chacun d'un groupe syntagmatique complet.

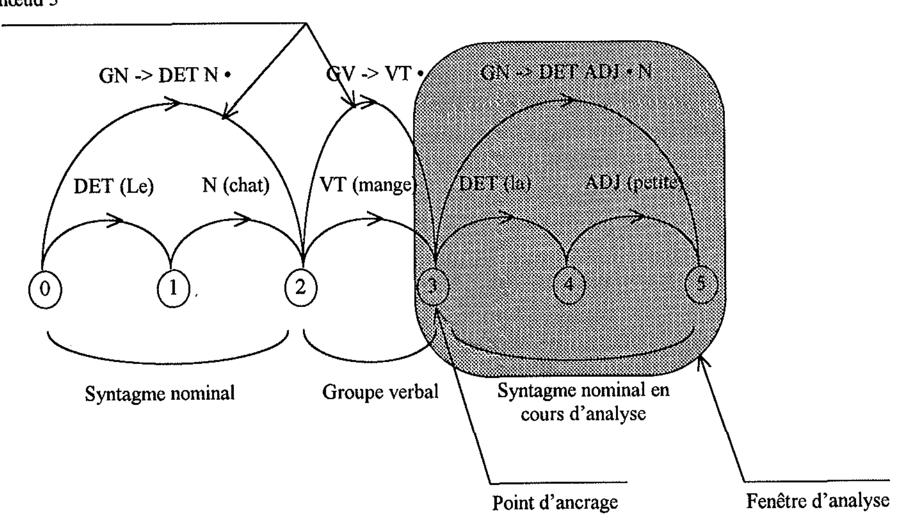


Figure 3: Exemple de segmentation syntagmatique en cours de réalisation. GN décrit un groupe nominal et GV,un groupe verbal. DET, N, VT, et ADJ sont les abréviations respectives de déterminant, nom, verbe transitif et adjectif.

Au commencement de l'analyse, le point d'ancrage correspond au nœud 0. Au fur et à mesure que les mots sont analysés, le module construit le ou les groupes syntaxiques partant de ce point d'ancrage. Quand le dernier mot reçu ne peut s'intégrer dans le groupe syntagmatique en cours d'analyse, on considère alors que ce mot appartient à un autre syntagme. On déplace alors le point d'ancrage vers la droite sur un nœud constituant la fin d'un groupe syntagmatique et on relance le processus. Si on ne trouve pas un tel nœud, on en déduit alors que ce dernier mot est incorrect et on le rejette. La segmentation

syntagmatique s'effectue par bonds successifs vers les extrémités des syntagmes déjà analysés. La segmentation appliquée à la phrase ambiguë

la belle ferme le voile

permet d'obtenir un treillis des groupes syntagmatiques présents dans la phrase (Figure 3).

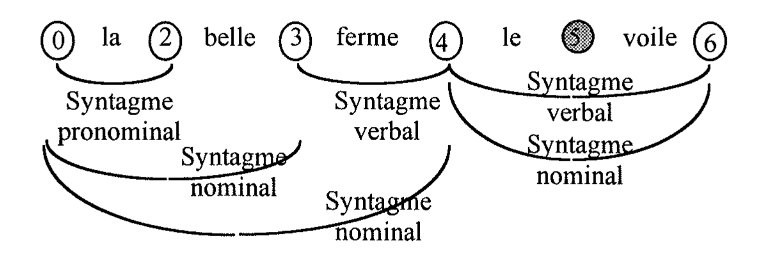


Figure 4: Treillis final des syntagmes présents dans la phrase

5 *Le module d'élagage*

Lorsqu'on considère le treillis de la Figure 4, on remarque au moins un syntagme dont la présence ne semble pas se justifier. En effet, l'interprétation du mot *la* en tant que syntagme pronominal ne peut être considérée comme acceptable pour deux raisons. En premier lieu, il n'existe pas de phrases affirmatives où un tel pronom est suivi par un adjectif; un tel syntagme est généralement placé entre le groupe sujet et le groupe verbal. De plus, même en considérant des formes syntaxiques ou stylistiques autorisant une telle construction, on remarque que ce syntagme s'incorpore mal à la phrase, car il n'est suivi par aucun autre syntagme. Il s'agit donc d'un résultat que l'on peut assurément mettre de côté du fait de son incompatibilité avec le reste de la phrase. Cette analyse de la cohérence d'une solution est réalisée par le *module d'élagage.*

Le travail effectué par ce module repose sur l'application d'un principe simple qui consiste à explorer le voisinage immédiat d'un syntagme pour vérifier s'il forme, avec ses voisins, une séquence acceptable. Cette technique s'apparente à celles utilisées dans les programmes d'étiquetage lexical (Brill, 1995), sauf que, dans notre cas, on se place à un niveau syntagmatique. Ce contrôle s'effectue par des règles de précédence qui indiquent les différents bigrammes syntagmatiques possibles (Figure 5).

Un syntagme n'est acceptable que si au moins l'un de ses prédécesseurs directs forme avec lui un bigramme syntagmatique correct. On applique aussi la réciproque de ce principe: un syntagme ne pouvant s'associer avec aucun de ses successeurs est considéré comme incorrect.

Séquences admissibles	Explications	Exemples
GN GV	Un groupe verbal peut suivre un groupe nominal	***Je vois** une voiture*
PRONOM_ANTE GV	Un pronom objet peut précéder un groupe verbal	*Je **la vois***
GN PRONOM_ANTE	Un pronom objet antéposé peut suivre un groupe nominal	***Je le** regarde*

Figure 5: Exemples de règles de précédence

On effectue donc un double balayage des séquences syntagmatiques pour que ce contrôle soit réalisé sur chacun des deux côtés de chaque syntagme. Les segments ne possédant pas de prédécesseurs ou de successeurs (à l'exception évidente des syntagmes de début et de fin de phrase) sont également écartés. À cause de ce manque de voisin, ils ne font manifestement pas partie d'une séquence syntagmatique reliant les deux extrémités de la phrase et ne peuvent donc s'intégrer à l'ensemble de celle-ci (Figure 6).

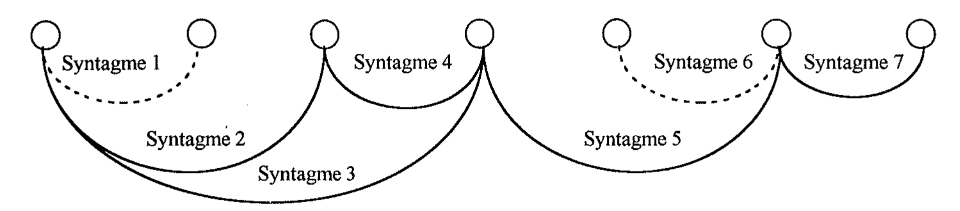


Figure 6: Treillis comportant deux syntagmes non intégrables au reste de la phrase par manque de voisins : les syntagmes 1 et 6.

La mise à l'écart d'un ou plusieurs arcs syntagmatiques peut entraîner que leurs voisins ne soient plus acceptables. Le programme effectue donc un balayage des différents segments de gauche à droite puis de droite à gauche jusqu'à ce que l'on ne constate plus d'exclusion.

6 *Le module d'analyse des dépendances*

Le *module d'analyse des dépendances* se charge d'accomplir la dernière étape de l'analyse à partir des éléments du niveau 2. Pour cela, il cherche à établir, au niveau 3 du Chart, des liaisons de dépendance entre les divers syntagmes retenus. À cause des diverses ambiguïtés qui peuvent persister à ce niveau de l'analyse, l'algorithme de dépendance se fonde sur un système de recherche en profondeur. À chaque étape de l'analyse, il choisit la dépendance ou le syntagme le plus probable. Ce choix repose sur l'application d'heuristiques topologiques. Quand un choix entre plusieurs possibilités se présente au cours de l'analyse, le système effectue un tri sur les différentes options qu'il peut suivre. Les heuristiques topologiques utilisées pour organiser ce tri sont les suivantes:

- Quand plusieurs syntagmes sont en concurrence au départ d'un même sommet, on privilégie le syntagme le plus long.
- Quand plusieurs dépendances sont en lice pour établir une tête – un régissant – sur un syntagme, on privilégie la dépendance pour laquelle la distance entre le syntagme à rattacher et sa tête est la plus courte.

L'application de ces heuristiques permet d'obtenir une liste ordonnée des différentes possibilités. Après cette phase d'ordonnancement, l'algorithme poursuit l'analyse en considérant l'hypothèse en tête de liste. En renouvelant ce processus sur tous les segments restants, on aboutit à une première analyse complète. On stocke alors la solution trouvée et on compte le nombre de syntagmes sans tête qu'elle contient. Ce nombre correspond au critère que nous avons retenu pour évaluer la qualité d'une analyse syntaxique: plus ce nombre est faible, meilleure est la fiabilité de l'analyse.

L'algorithme revient ensuite sur le dernier choix laissé en suspens et explore une nouvelle analyse en considérant l'hypothèse suivante. Si, au cours d'une telle exploration, on s'aperçoit que le nombre de syntagmes sans tête établis par ce parcours dépasse celui de la meilleure solution, on abandonne cette analyse. Si, par contre, une exploration débouche sur une solution équivalente voire meilleure que la précédente, on stocke cette nouvelle solution et on met à jour le score établi précédemment. De cette manière, on restreint l'espace de recherche en ne s'attachant qu'aux solutions jugées les meilleures à un instant donné de l'analyse. La Figure 6 présente une illustration du résultat généré par l'analyseur pour une des phrases du corpus.

7 *Implantation, résultats et applications*

Le système décrit a été implanté en C++ en tirant parti de la programmation orientée objet. Il a été développé avec la perspective d'une utilisation sous divers environnements parmi lesquelles Windows et Unix. La version sous Windows bénéficie d'une interface graphique permettant d'étudier et d'améliorer dans de bonnes conditions le fonctionnement du programme.

L'analyseur syntagmatique présent dans le module de segmentation a été intégré au sein d'un système de dialogue oral. Cette application consiste en la mise en œuvre d'un agent conversationnel de navigation permettant à un utilisateur de se déplacer à la voix au sein d'un environnement de réalité virtuelle (Godéreaux et al., 1996). L'analyseur réalise une analyse purement formelle permettant de couvrir la quasi totalité des ordres de déplacement du corpus que nous avons collecté pour cette application.

Nous avons également évalué notre système sur un échantillon de 54 phrases issues d'un corpus de comptes rendus médicaux. La combinaison des deux premiers modules a donné de très bons résultats. Le module de segmentation a détecté plus de 93% des segments attendus. Une seule phrase de l'échantillon n'a pu être segmentée totalement par l'analyseur à cause d'une incise entre une préposition et son groupe nominal. Sans ce cas

particulier, le taux de couverture des segments correctement détectés aurait atteint plus de 99,5%.

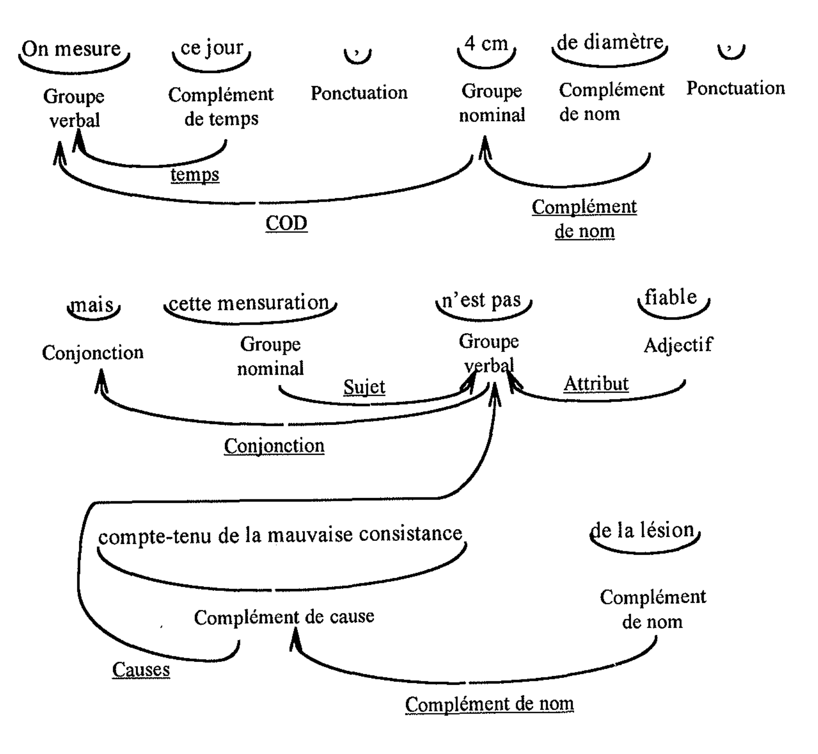


Figure 7: Exemple de mise en relation des segments pour une des phrases du corpus

Ces résultats sont cependant accompagnés d'une surproduction de segments importante avec plus de 191% de segments incorrects par rapport au nombre de segments attendus. Le module d'élagage apporte un gain incontestable au niveau de la précision du système en éliminant plus de 82% de ces segments incorrects, faisant ainsi tomber le taux de surproduction aux alentours de 33%. Les résultats obtenus par le module de dépendance portent sur les phrases ayant passé avec succès les deux premières phases de l'analyse. Nous avons donc exclus de l'échantillon la phrase qui avait posé quelques difficultés au module de segmentation à cause d'une incise rejetée par la grammaire de l'analyseur. Certaines des 53 phrases restantes pouvant être interprétées de multiples manières, un ensemble de 107 analyses plausibles d'un point de vue sémantique a été établi comme modèle d'analyse de l'échantillon. Sur ces 107 solutions, l'analyseur en a trouvé 88 sans aucune erreur (82,2 %). Si on fait abstraction des erreurs mineures de

catégorisation ou de typage des segments ou des relations, ce taux de réussite monte à 83,2 %.

8 *Conclusion*

Nous avons présenté un analyseur syntaxique fondé sur les techniques de Chart. L'analyse proprement dite se décompose en trois phases. Grâce à ces trois phases de traitement, le système peut construire une analyse progressive. Celle-ci débute au niveau lexical, se poursuit au niveau syntagmatique et se termine au niveau de l'organisation structurelle de la phrase elle-même. Cette décomposition permet d'accomplir une analyse plus rapide et robuste.

L'organisation à plusieurs niveaux du Chart favorise l'analyse partielle et devrait permettre la mise en œuvre d'algorithmes de reprise sur erreur pour traiter les phrases agrammaticales. Enfin, nous avons confronté un sous-ensemble de notre système à la navigation orale dans les mondes virtuels ainsi qu'à un échantillon de phrases issues d'un corpus de comptes rendus médicaux.

Les résultats que nous avons obtenus nous permettent de valider notre démarche et d'envisager une analyse complète ou partielle de rapports médicaux sans restriction. Une telle analyse devrait permettre notamment la mise en œuvre d'outils perfectionnés de recherche d'informations et de compréhension automatique de textes.

Références

Brill E.: *Transformation-Based Error-Driven Learning and Natural Language Processing:* A Case Study in Part of Speech Tagging, *Computational Linguistics*, 21(4):543-565, 1995.

Chanod, J.P. & Tapanainen, P.: *Tagging French - Comparing a Statistical and Constrainst-Based Method,* Proceeding of the Seventh Conference of the European Chapter of the ACL (EACL'95), page 149-156, Association for Computational Linguistics, Dublin, 1995.

Chomsky, N.: *Syntactic Structures*, La Haye, Mouton, 1957.

Covington, M.: *Parsing discontinuous constituents in dependency grammar*, Computational Linguistics 16:234-236, 1990.

Earley, J.: *An efficient context-free parsing algorithm*. Communication of the ACM 13:94-102, 1970.

El Guedj, P.O. & Nugues, P.: *A Chart parser to analyze large medical corpora,* In: Proceedings of the 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Baltimore, pp. 1404-1405, November 1994.

El Guedj P.-O.: *Analyse syntaxique par chart combinant règles de dépendance et règles syntagmatiques*, Mémoire de thèse, Université de Caen, 10 septembre 1996.

El Guedj P.-O. & P. Nugues: *Un analyseur syntaxique combinant plusieurs formalismes au sein d'un chart hiérarchique*, Actes du Colloque RÉCITAL '96, pp. 83-91, Dourdan, septembre 1996.

Gazdar, G. & Mellish, C.: *Natural language Processing in Prolog: an introduction to computational linguistics.* Wokingham: Addison-Wesley, 1989.

Godéreaux, C., Diebel, K., El-Guedj, P.O., Revolta, F. & Nugues, P.: *An Interactive Spoken Dialog Interface to Virtual Worlds,* in: Linguistic Concepts and Methods in CSCW, J.H. Connolly & L. Pemberton eds, Chapter 13, pp. 177-200, Springer Verlag, 1996.

Kaplan, R.M.:, *A general syntactic processor*, In natural Language Processing, pp. 193-241, Randall Rustin (ed.), Algorithmics Press, New York, 1973.

Kay, M.: *The MIND* system, In Natural Language Processing, pp. 155-188, Randall Rustin (ed.), Algorithmics Press, New York, 1973.

Mel'cuk I. A.: *Dependency Syntax: Theory and Practice*, State University of New York Press, 1988.

Tesnière, L.: *Éléments de la syntaxe structurale*, Paris, Klincksieck, 1959.
Wirén, M.: *Studies in Incremental Natural-Language Analysis*, Dissertation N°. 292, Linköping Studies in Science and Technology, University of Linköping, Suède, 1992.

Reçu le 18-07-1997

Adresse des auteurs: Pierre-Olivier El Guedj et Pierre Nugues, ISMRA et Univeristé Caen, 6, Bd du Maréchal Juin, F-14050 Caen

A Parser Combining two Syntactic Formalisms (Summary)
In this paper, we describe a parser combining a phrase structure grammar, a set of precedence rules and a dependency grammar. The system is based on a Chart split in several hierarchical levels. These levels correspond to different stages of the parsing process managed by three modules. Each of these modules collaborate to the parsing bringing its linguistic competence to the system. The *Segmentation Module* carries out a segmentation of the sentence to give a lattice of the constituents (Subject, Verb, Adverbial Complements, etc.), structural words (conjunctions, ...) and punctuation signs. The *Pruning Module* carries out a selection among the lattice elements in order to retain only the segments which are well integrated with the rest of the sentence. At last, the *Dependency Parser Module* links these segments to construct the dependency graph of the sentence.

Sintaksa analizo kombinanta du formalismojn ene de la hierarĥia ĉarto (resumo)
En tiu ĉi artikolo ni priskribas analizilon kombinantan frazstrukturan gramatikon, aron de precedentaj leĝoj kaj la dependan gramatikon. La sistemo baziĝas sur la ĉarto dispecigita en plurajn hierarkiajn nivelojn. Tiuj ĉi niveloj korespondas al diversaj etapoj de la analiziga procezo direktita pere de tri moduloj. Ĉiu el la moduloj kunlaboras kun la analizado, alportante sian lingvistikan kompetentecon al la sistemo. La *segmentiga modulo* alportas segmentigon de la frazo por doni kradaron de konsisteroj (subjekto, verbo, adverbaj komplementoj ktp.), stukturajn vortojn (konjunkcioj...) kaj interpunkciajn signojn. La *foriga modulo* alportas elekton inter la kradaj elementoj por konservi nur tiujn segmentojn, kiuj bone integriĝas kun la resto de la frazo. Laste, la *dependa analiza modulo* kunligas tiujn segmentojn por konstrui la dependan grafikaĵon de la frazo.

Richtlinien für die Manuskriptabfassung

Artikel von mehr als 12 Druckseiten Umfang (ca. 36.000 Anschläge) können in der Regel nicht angenommen werden; bevorzugt werden Beiträge von maximal 8 Druckseiten Länge. Außer deutschsprachigen Texten erscheinen ab 1982 regelmäßig auch Artikel in den drei Kongreßsprachen der Association Internationale de Cybernétique, also in Englisch, Französisch und Internacia Lingvo. Die verwendete Literatur ist, nach Autorennamen alphabetisch geordnet, in einem Schrifttumsverzeichnis am Schluß des Beitrags zusammenzustellen - verschiedene Werke desselben Autors chronologisch geordnet, bei Arbeiten aus demselben Jahr nach Zufügung von „a", „b" usf.. Die Vornamen der Autoren sind mindestens abgekürzt zu nennen. Bei selbständigen Veröffentlichungen sind anschließend nacheinander Titel (evt. mit zugefügter Übersetzung, falls er nicht in einer der Sprachen dieser Zeitschrift steht), Erscheinungsort und -jahr, womöglich auch Verlag, anzugeben. Zeitschriftenbeiträge werden nach dem Titel vermerkt durch Name der Zeitschrift, Band, Seiten und Jahr. - Im Text selbst soll grundsätzlich durch Nennung des Autorennamens und des Erscheinungsjahrs (evt. mit dem Zusatz „a" etc.) zitiert werden. - Bilder (die möglichst als Druckvorlagen beizufügen sind) einschl. Tabellen sind als „Bild 1" usf. zu numerieren und nur so zu erwähnen, nicht durch Wendungen wie „vgl. folgendes (nebenstehendes) Bild". - Bei Formeln sind die Variablen und die richtige Stellung kleiner Zusatzzeichen (z.B. Indices) zu kennzeichnen.
Ein Knapptext (500 - 1.500 Anschläge einschl. Titelübersetzung) ist in mindestens einer der drei anderen Sprachen der GrKG/Humankybernetik beizufügen.
Im Interesse erträglicher Redaktions- und Produktionskosten bei Wahrung einer guten typographischen und stilistischen Qualität ist von Fußnoten, unnötigen Wiederholungen von Variablensymbolen und übermäßig vielen oder typographisch unnötig komplizierten Formeln (soweit sie nicht als druckfertige Bilder geliefert werden) abzusehen, und die englische oder französische Sprache für Originalarbeiten in der Regel nur von „native speakers" dieser Sprachen zu benutzen.

Direktivoj por la pretigo de manuskriptoj

Artikoloj, kies amplekso superas 12 prespaĝojn (ĉ. 36.000 tajpsignojn) normale ne estas akceptataj; preferataj estas artikoloj maksimume 8 prespaĝojn ampleksaj. Krom germanlingvaj tekstoj aperadas de 1982 ankaŭ artikoloj en la tri kongreslingvoj de l'Association Internationale de Cybernétique, t.e. en la angla, franca kaj Internacia lingvoj.
La uzita literaturo estu surlistigita je la fino de la teksto laŭ aŭtornomoj ordigita alfabete; plurajn publikaĵojn de la sama aŭtoro bv. surlistigi en kronologia ordo, en kazo de samjareco aldoninte „a", „b" ktp.. La nompartoj ne ĉefaj estu almenaŭ mallongigite aldonitaj. De disaj publikaĵoj estu - poste - indikitaj laŭvice la titolo (evt. kun traduko, se ĝi ne estas en unu el la lingvoj de ĉi tiu revuo), la loko kaj jaro de la apero, kaj laŭeble la eldonejo. Artikoloj en revuoj ktp. estu registritaj post la titolo per la nomo de la revuo, volumo, paĝoj kaj jaro. – En la teksto mem bv. citi pere de la aŭtornomo kaj la aperjaro (evt. aldoninte „a" ktp.). – Bildojn (laŭeble presprete aldonendajn!) inkl. tabelojn bv. numeri per „bildo 1" ktp. kaj mencii ilin nur tiel, neniam per teksteroj kiel „vd. la jenan (apudan) bildon". – En formuloj bv. indiki la variablojn kaj la ĝustan pozicion de etliteraj aldonsignoj (ekz. indicoj).
Bv. aldoni resumon (500 –1.500 tajpsignojn inkluzive tradukon de la titolo) en unu el la tri aliaj lingvoj de GrKG/Humankybernetik.
Por ke la kostoj de la redaktado kaj produktado restu raciaj kaj tamen la revuo grafike kaj stile bonkvalita, piednotoj, nenecesaj ripetoj de simboloj por variabloj kaj tro abundaj, tipografie nenecese komplikaj formuloj (se ne temas pri prespretaj bildoj) estas evitendaj, kaj artikoloj en la angla aŭ franca lingvoj normale verkendaj de denaskaj parolantoj de tiuj ĉi lingvoj.

Regulations concerning the preparation of manuscripts

Articles occupying more than 12 printed pages (ca. 36,000 type-strokes) will not normally be accepted; a maximum of 8 printed pages is preferable. From 1982 onwards articles in the three working-languages of the Association Internationale de Cybernétique, namely English, French and Internacia Lingvo will appear in addition to those in German. Literature quoted should be listed at the end of the article in alphabetical order of authors' names. Various works by the same author should appear in chronological order of publication. Several items appearing in the same year should be differentiated by the addition of the letters "a", "b", etc. Given names of authors, (abbreviated if necessary, should be indicated. Works by a single author should be named along with place and year of publication and publisher if known. If articles appearing in journals are quoted, the name, volume, year and page-number should be indicated. Titles in languages other than those of this journal should be accompanied by a translation into one of these if possible. – Quotations within articles must name the author and the year of publication (with an additional letter of the alphabet if necessary). – Illustrations (fit for printing if possible) should be numbered "figure 1", "figure 2", etc. They should be referred to as such in the text and not as, say, "the following figure". – Any variables or indices occuring in mathematical formulae should be properly indicated as such.
A resumee (500 – 1,500 type-strokes including translation of title) in at least one of the other languages of publication should also be submitted.
To keep editing and printing costs at a tolerable level while maintaining a suitable typographic quality, we request you to avoid footnotes, unnecessary repetition of variable-symbols or typographically complicated formulae (these may of course be submitted in a state suitable for printing). Non-native-speakers of English or French should, as far as possible, avoid submitting contributions in these two languages.

Forme des manuscrits

D'une manière générale, les manuscrits comportant plus de 12 pages imprimées (env. 36.000 frappes) ne peuvent être acceptés; la préférence va aux articles d'un maximum de 8 pages imprimées. En dehors de textes en langue allemande, des articles seront publiés régulièrement à partir de 1982, dans les trois langues de congrès de l'Association Internationale de Cybernétique, donc en anglais, français et Internacia Lingvo.
Les références litteraires doivent faire l'objet d'une bibliographie alphabétique en fin d'article. Plusieurs œuvres d'un même auteur peuvent être énumérées par ordre chronologique. Pour les ouvrages d'une même année, mentionnez "a", "b" etc. Les prénoms des auteurs sont à indiquer, au moins abrégés. En cas de publications indépendantes indiquez successivement le titre (eventuellement avec traduction au cas où il ne serait pas dans l'une des langues de cette revue), lieu et année de parution, si possible éditeur. En cas d'articles publiés dans une revue, mentionnez après le titre le nom de la revue, le volume/tome, pages et année. – Dans le texte lui-même, le nom de l'auteur et l'année de publication sont à citer par principe (eventuellement complétez par "a" etc.). – Les illustrations (si possible prêtes à l'impression) et tables doivent être numérotées selon "fig. 1" etc. et mentionées seulement sous cette forme (et non par "fig. suivante ou ci-contre").
En cas de formules, désignez les variables et la position adéquate par des petits signes supplémentaires (p. ex. indices).
Un résumé (500–1.500 frappes y compris traduction du titre est à joindre rédigé dans au moins une des trois autres langues de la grkg/Humankybernetik.
En vue de maintenir les frais de rédaction et de production dans une limite acceptable, tout en garantissant la qualité de typographie et de style, nous vous prions de vous abstenir de bas de pages, de répétitions inutiles de symboles de variables et de tout surcroît de formules compliquées (tant qu'il ne s'agit pas de figures prêtes à l'impression) et pour les ouvrages originaux en langue anglaise ou en langue française, recourir seulement au concours de natifs du pays.